PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-126070

(43)Date of publication of application : 11.05.2001

(51)Int.Cl. G06T 7/00
G06T 1/00
H04N 1/387
H04N 5/232
H04N 5/262

(21)Application number : 11-306884

(22)Date of filing : 28.10.1999

(71)Applicant : ATR MEDIA INTEGRATION & COMMUNICATIONS RES LAB

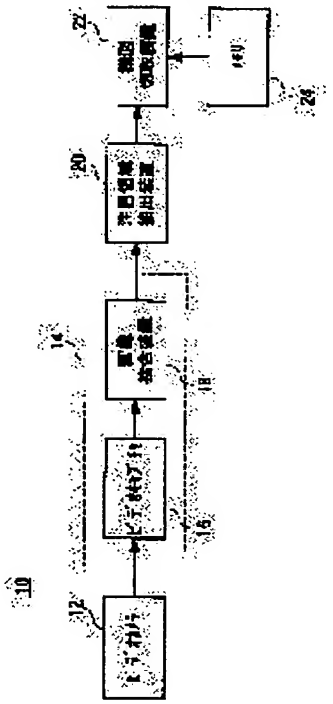
(72)Inventor : TANAKA SHOJI
IWADATE YUICHI

(54) NOTICING AREA EXTRACTING DEVICE AND AUTOMATIC COMPOSITION DECIDING DEVICE

(57)Abstract:

PROBLEM TO BE SOLVED: To extract a noticing area adjusted to human subjectively and to automatically device well-balanced composition.

SOLUTION: A noticing area extracting device and an automatic composition deciding device 10 include an image forming device 14, which generates the original image of a panoramic image from video photographed by a camera 12. A noticing area extracting device 20 extracts a noticing area from the original image given from the device 14. Namely, evaluation matched with human subjectivity is given in accordance with the physical feature of the original image and the noticing area is extracted in accordance with the evaluated result. A composition cutting off device 22 cuts out the extracted noticing area and an adjacent image area from the original image by referring to data on paintings painted by painters and photographs taken by photographers stored in a memory 24. Namely, the data can be cut out by the same composition as the case with painted images or photographed images.



LEGAL STATUS

[Date of request for examination] 17.01.2001

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of

* NOTICES *

1.This document has been translated by computer. So the translation may not reflect the original precisely.
2.**** shows the word which can not be translated.
3.In the drawings, any words are not translated.

CLAIMS

[Claim(s)]
[Claim 1] An attention field extractor equipped with an evaluation means to be the attention field extractor which extracts an attention field from a subject-copy image, and to evaluate the degree of [**] based on the physical feature, and an extraction means to extract the aforementioned attention field according to the evaluation result of the aforementioned evaluation means.
[Claim 2] The aforementioned physical feature is an attention field extractor containing the degree of different kind of a color according to claim 1.
[Claim 3] For the aforementioned evaluation means, the aforementioned physical feature is an attention field extractor according to claim 2 by which the degree of [******] is evaluated based on at least one degree of different kind among the four degrees of different kind, including further the degree of different kind of the formal degree of different kind, and area, and the degree of different kind of a texture.
[Claim 4] The aforementioned physical feature is an attention field extractor according to claim 1 to 3 which contains a color further.
[Claim 5] The aforementioned physical feature is an attention field extractor according to claim 4 which contains further the area and spatial frequency of a field in the aforementioned subject-copy image.
[Claim 6] The attention field extractor according to claim 1 to 5 further equipped with a photography means to photo a desired image, and a picture generation means to generate the aforementioned subject-copy image based on the aforementioned image.
[Claim 7] The aforementioned picture generation means is an attention field extractor including a junction means to join the aforementioned image photoed with the aforementioned camera for every frame according to claim 6, including the driving means to which the aforementioned photography means carries out the rotation drive of a camera and the aforementioned camera.
[Claim 8] Automatic composition determination equipment equipped with a maintenance means to hold the data about the criteria picture which is automatic composition determination equipment using the attention field extractor according to claim 1 to 7, and has criteria composition, and the cutoff means which cuts out the picture of an attention field from a subject-copy image with reference to the aforementioned criteria composition.
[Claim 9] The aforementioned maintenance means is automatic composition determination equipment according to claim 8 which holds the data about two or more aforementioned criteria pictures, and is further equipped with a selection means to choose the aforementioned data which suited the picture of the aforementioned attention field.
[Claim 10] The aforementioned data are automatic composition determination equipment according to claim 8 or 9 which contains the image data corresponding to the aforementioned criteria picture, a photographic subject's configuration data, and a photographic subject's position data at least.

[Translation done.]

* NOTICES *

DETAILED DESCRIPTION

[Detailed Description of the Invention]
[0001]
[Industrial Application] This invention relates to the attention field extractor cut out from a subject-copy image so that the composition of a picture that experts, such as a painter and a photographer, made the attention field which extracted and extracted the attention field may be suited and it may be especially settled, for example from a subject-copy image about the automatic composition determination equipment which used an attention field extractor and it, and the automatic composition determination equipment using it.
[0002] Here, the field where people observe an attention field in a picture or an image is said.
[0003]
[Description of the Prior Art] As this kind of conventional technology, various methods are proposed about the technique to which a watcher extracts an attention field from a picture. (1) Milanese and Itti And Koch and others assumes the discontinuous part in a picture to be an attention field, and generates the feature map (shade picture) corresponding to two or more physical features acquired from a picture, next asks for the discontinuous part of each feature map, and is extracting as an attention field what unified them.
[0004] (2) Milanese and others filtered the feature map using two or more Difference-of-ofiented-Gaussians filters with which the size was fixed in the discontinuous part of each feature map, and chose and searched for the filtering result from which an output serves as the maximum. Itti, Koch and others normalized each feature map with the square error with each average, and unified all the feature maps by linear combination. And the integrated feature map was recursively filtered with the Difference of Gaussian filter, and the local peak of the filtering result finally obtained was extracted as an attention field.
[0005] Thus, Milanese, Itti and others were extracting the attention field by processing of pixel level, such as filtering and relaxation.
[0006] (3) Martin, Takeuchi and others evaluated the brightness information obtained from a picture based on the information theory of Shannon, and made the attention field the portion with the high amount of information acquired as a result. By this method, the field, i.e., the field which looks complicated, where distribution is large and the bright field of a brightness value are mainly extracted.
[0007] Moreover, with the conventional camera, human being had determined the composition of a photographic subject and a photographic subject manually.
[0008]
[Problem(s) to be Solved by the Invention] However, since the size of an attention field changed with pictures by the technique of of (1) and (2), it was difficult to extract an attention field exactly using the filter with which the size was fixed. Moreover, in the case of the picture which cannot not necessarily be said that the field which looks complicated, or a bright field is in agreement with an attention field like the black picture of a flower vase placed, for example in front of the wall of a complicated pattern by the technique of (3), it was difficult to extract an attention field exactly. Furthermore, by such proposal, there were few examples which performed collating with a watcher's (human being) subjectivity and an extraction result, and it was a question whether the attention field which actually suits human being's subjectivity can be extracted.
[0009] Moreover, with the conventional camera, when ordinary men without the sense over a photograph determine composition, the good photograph of balance cannot necessarily be taken.
[0010] So, the main purpose of this invention is offering the attention field extractor which can extract exactly the attention field which suited human being's subjectivity.
[0011] Moreover, other purposes of this invention are offering the automatic composition determination equipment which is automatic and can determine the good composition of balance.

[0012]
[Means for Solving the Problem] The 1st invention is an attention field extractor which extracts an attention field from a subject-copy image, and is an attention field extractor equipped with an evaluation means to evaluate the degree of [**] based on the physical feature (whenever for it to be conspicuous), and an extraction means to extract an attention field according to the evaluation result of an evaluation means.
[0013] The 2nd invention is automatic composition determination equipment which used the attention field extractor of a publication for the 1st invention, and is automatic composition determination equipment equipped with a maintenance means to hold the data about the criteria picture which has criteria composition, and the cutoff means which cuts out the picture of an attention field from a subject-copy image with reference to criteria composition.
[0014]
[Function] An evaluation means evaluates the degree of [**] by the attention field extractor of the 1st invention according to the physical feature of a subject-copy image. Here, the degree of [**] means the parameter suitable for human being's subjectivity. An extraction means extracts a most conspicuous field from an evaluation result as an attention field. That is, since an evaluation means carries out evaluation which suited human being's subjectivity according to the physical feature, the attention field which suited human being's subjectivity can be extracted.
[0015] For example, when the physical feature contains the degree of different kind of a color, the degree of [**] can be evaluated based on the difference in the color of each field.
[0016] Moreover, since it contains further the degree of different kind of the formal degree of different kind, and area, and the degree of different kind of a texture (pattern) whenever the color of the physical feature is heterogeneous in addition, if the degree of [**] is evaluated based on at least one degree of different kind of these four degrees of different kind, according to the feature of a subject-copy image, the degree of [**] can be evaluated exactly.
[0017] Moreover, if it is the case where it evaluates also about three elements (a hue, saturation, lightness) of a color, the field near the conspicuous color (red) by human being's subjectivity can be estimated as a most conspicuous field.
[0018] Furthermore, if it evaluates also about the area of each field in spatial frequency or a subject-copy image, most conspicuous evaluation of a field can be judged still more exactly.
[0019] Moreover, the picture corresponding to the image which photoed the desired image with the photography means, for example, was photoed can be compounded, and a subject-copy image can also be generated.
[0020] For example, if the position and height of a camera which are contained in a photography means are fixed, 360 degrees rotates and it enables it to take a photograph, when a junction means joins the picture in every frame, the subject-copy image of a panorama picture is generable within the limits of 360 degrees.
[0021] With the automatic composition determination equipment of the 2nd invention, a maintenance means makes a criteria picture the pictures picture and photograph corresponding to the photograph which the pictures which the painter drew, and the photographer took, and the data about this criteria picture are held. Since a cutoff means cuts out the picture of an attention field from a subject-copy image with reference to the data about this criteria picture, it is automatic and can determine the good composition of balance.
[0022] Moreover, if the data about two or more criteria pictures are held, since a selection means can choose the composition suitable for the picture of an attention field, the good composition of balance can be determined about all pictures.
[0023] Since they contain the image data corresponding to a criteria picture, a photographic subject's configuration data, and a photographic subject's position data at least, the data about the above criteria pictures choose the composition suitable for the attention field, are automatic and can determine the good composition of balance. That is, a photograph as if it photoed the photographic subject in the composition is generable.
[0024]
[Effect of the Invention] Since an attention field is extracted according to the evaluation result of the physical feature suitable for human being's subjectivity according to this invention, the attention field which suited human being's subjectivity can be extracted.
[0025] According to other invention, since an attention field is cut out from a subject-copy image using the criteria composition of a criteria picture, it is automatic and the good composition of balance can be determined.
[0026] The above-mentioned purpose of this invention, the other purposes, the feature, and an advantage will become still clearer from the detailed explanation of the following examples given with reference to a drawing.
[0027]
[Example] With reference to drawing 1 , the automatic composition determination equipment 10 of this example contains a video camera (only henceforth a "camera") 12. A photography position (a position and height) is fixed using a tripod etc., 360 degrees of cameras 12 rotate and they can be photoed. The image photoed with the camera 12 is inputted into picture generation equipment 14, and the panorama picture as a subject-copy image is generated from the

photoed image. In addition, since 360 degrees of cameras 12 can be rotated, picture generation equipment 14 can generate a panorama picture within the limits of 360 degrees.

[0028] Picture generation equipment 14 generates a panorama picture according to the flow view shown in drawing 2 including a video capture 16 and picture junction equipment 18. That is, if photography is started with a camera 12, after picture generation equipment 14 starts processing, is Step S1, carries out the capture of the first video frame (the 1st frame) using a video capture 16 and generates a synthetic picture from the picture corresponding to oneth of them, it will generate a gray-scale picture from the synthetic picture using picture junction equipment 18.

[0029] At continuing Step S3, the capture of the following frame is carried out and a gray-scale picture is generated like the 1st frame. At Step S5, in order to determine from which position the synthetic picture corresponding to the present frame (object frame) should be compounded, a search template is started from the gray-scale picture of an object frame as shown in drawing 3 . The width of face of this search template is 50 pixels, and the height is the same as a frame picture. Moreover, this search plate is cut down from the zero of an object frame.

[0030] Then, at Step S7, the search range is set up from the gray-scale picture of a synthetic picture, and the picture of the same size as a search template is started from the search range. That is, the width of face of the search range is 100 pixels, and the height is the same as a frame picture. Moreover, when the upper left of a frame picture is made into a zero (x y) (= (1 1)), the x-coordinate of a search plate is the position which subtracted only 100 from the width of face of a frame picture, and the y-coordinate is 1.

[0031] Next, at Step S9, from the search range, the picture of the same size as a search template is started, and the absolute value of the difference of the mutual pixel value corresponding to the picture and search plate which were cut down is computed. At continuing Step S11, it judges whether the difference of a pixel value is minimum (0). If it is "YES" at Step S11, the started picture will judge that it is the same size as a search plate, and will progress to Step S13. On the other hand, if it is "NO" at Step S11, after judging that the sizes of the picture and search plate which were cut down differ and shifting the search range of 1 pixel to the positive direction of a x axis at Step S15, it will return to Step S5. Thus, processing is repeated until the picture and search template which it started 1 pixel at a time with staggering become the same size.

[0032] At Step S13, an X coordinate in case the absolute value of a difference is minimum (0) is computed. At continuing Step S17, a synthetic picture and an object frame picture are piled up by the computed X coordinate, and a new synthetic picture is generated. And at Step S19, it judges whether the gray-scale picture was generated from the generated synthetic picture, and the panorama picture was generated at Step S21. If the panorama picture will not be generated if it is "NO" at Step S21 that is, it returns to Step S3. On the other hand, if it is "YES" at Step S21, it will judge that the panorama picture (subject-copy image) as shown in drawing 4 (A) was generated, a subject-copy image will be outputted to the latter attention field extractor 20 (drawing 1) at Step S23, and processing will be ended.

[0033] Returning to drawing 1 , the attention field extractor 20 extracts a most conspicuous field (attention field) from the given subject-copy image. Specifically, the attention field extractor 20 processes according to the flow view shown in drawing 5 . That is, if a subject-copy image is given from picture generation equipment 14, the attention field extractor 20 will start processing and will carry out field division of the subject-copy image at Step S31. Specifically, as shown in drawing 4 (B), a subject-copy image is divided into a drawing field and a picture field. the method of this field division -- 1997IEEE -- setting -- W.Y.Ma and B.S.Manjunath ** -- the boundary detection method based on "edge flow" indicated to "Edge Flow:A Framework of Boundary Detection and Image Segmentation" is applied If it explains briefly, this method is edge flow which asks for the change direction of a color or a pattern in each place of a picture, and consists of strength (size) of the change direction and change. A vector is determined. And it is edge flow by repetitive operation. A vector is spread in the direction of each vector, and let the place where final vectors collide be the boundary line of each field.

[0034] Therefore, the divided drawing field as shown in drawing 4 (C) is extracted, and Step S33 estimates the degree of a drawing field of [**] at Step S35. That is, it asks for the degree parameter of a drawing field of [**]. Here, the subjectivity evaluation experiment which the artificer conducted showed that the physical features required for evaluation of the degree of [**] were the degree of different kind of the degree of different kind of the degree of different kind of a color, and a texture, and a form and the degree of different kind of area, a color, spatial frequency, and area. Moreover, generally, since the relation between human being's subjectivity evaluation result and a physical characteristic is expressed with a serpentine curve in many cases, an one ****** evening function is used for evaluation which is the degree of [**] several 1.

[0035]

[Equation 1]

$$H(m,n) = \int_0^1 h^{m-1}(1-h^{n-1})\,dh$$

[0036] The performance index of the degree of [**] as shown in several 2 is defined using this beta function.
[0037]
[Equation 2]
$$A_{i,t} = \int_0^1 a_{i,t}^{m-1}(1-a_{i,t}^{n-1})\,da_{i,t}$$

$$a_i = HET_i + FP_{i,t}$$

ただし、

$A_{i,t}$: 領域iの時間tにおける誘目度
$a_{i,t}$: 領域iの時間tにおける特徴量統合値
HET_i : 領域iの異質性誘目度
$FP_{i,t}$: 領域iの時間tにおける特徴誘目度
i : 領域i
t : 時間
m, n : 立ち上がり, 飽和ﾊﾟﾗﾒｰﾀ

[0038] Moreover, the degree HET to the above-mentioned four degrees of different kind of each field shown in several 2 of [**] is defined by several 3.
[0039]
[Equation 3]
$$HET_i = wh_1 \cdot HC_i + wh_2 \cdot HT_i + wh_3 \cdot HS_i + wh_4 \cdot HSh_i$$

ただし、

HET_i : 領域iの異質性誘目度
HC_i : 領域iの色の異質度
HT_i : 領域iのﾃｸｽﾁｬの異質度
HS_i : 領域iの面積の異質度
HSh_i : 領域iの形の異質度
wh_i : 重み係数

[0040] Furthermore, the degree H of different kind of each physical feature is computed according to several 4, when standard deviation of dm and Difference d is set [the difference of the feature value and the average feature value of all fields] to std for the average of d and Difference d.
[0041]
[Equation 4]
$$H(d, dm, std) = \left|\frac{d-dm}{std}\right|$$

[0042] That is, the degree HC of different kind of the color of each field is CIE. L* a* b* The result which computed and computed the standard deviation of the color difference of the average color of a field and the average color of all fields, a color difference average, and the color difference using the color difference formula in a consciousness equal color space is substituted for several 4, and it asks. In addition, the color difference formula is indicated in detail by 1994 "a color science handbook." Specifically, it is the sexual desire news (R, G, B) of each pixel of a picture Equal perceived-color-space L* a* b* It changes and Euclidean distance in the space is made into the color difference. That is, a color difference formula is shown like several 5.
[0043]
[Equation 5]
$$\Delta Eab^* = [(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)]^{\frac{1}{2}}$$

[0044] Next, the degree HT of different kind of a texture is explained. a texture -- 1996IEEE -- setting -- B.S.Manjunath W.Y.Ma ** -- it is expressed by the texture feature vector indicated to "Texture Features for Browsing and Retrieval of Image Data", and the difference between textures is expressed by the Euclidean distance between texture feature vectors

[0045] Here, a texture feature vector is expressed by the vector which uses as an element the response when filtering a picture by the Gabor filter bank constituted from two or more Gabor filters with which a size differs from a direction. However, since there is no orthogonality in each response when it filters with two or more Gabor filters, redundant information may be included in the filtered result. Then, the parameter of each filter in the Gabor filter bank is determined by the technique indicated by above-mentioned 1996IEEE. Specifically, it asks for theta, a, sigmau, and sigmav of the Gabor filter shown by several 6 by several 7. In addition, for this technique, the filter which adjoins as shown in drawing 6 is Half-Peak. In order to determine that the scale (size) and direction parameter of a filter touch and to express a textural facility, 24 filters (four scales, six directions) are used.

[0046]

[Equation 6]

$$G_{mn}(x, y) = a^{-m} \cdot G(x', y')$$

$$G(x, y) = \left(\frac{1}{2\pi\sigma_x\sigma_y}\right) \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right] \cdot \exp(2\pi jUx)$$

$$U = a^{m}$$

$$x' = x\cos\theta + y\sin\theta$$

$$y' = -x\sin\theta + y\cos\theta$$

$$\sigma_x = \frac{1}{2\pi\sigma_u}$$

$$\sigma_y = \frac{1}{2\pi\sigma_v}$$

ただし、

Gmn(x, y) : スケールID=m, 方位ID=nの時のガボールフィルタ
m : 0 .. S-1
n : 0 .. K-1
S : スケール数
K : 方位数

[0047]

[Equation 7]

$$\theta = \frac{n\pi}{K}$$

$$a = (U_h/U_l)^{\frac{1}{K-1}}$$

$$\sigma_u = \frac{(a-1)U}{(a+1)\sqrt{2\ln 2}}$$

$$\sigma_r = \tan\left(\frac{\pi}{2k}\right)\left[U_h - 2\ln 2\left(\frac{\sigma_u^2}{U_h}\right)\right]\left[2\ln 2 - \frac{(2\ln 2)^2\sigma_u^2}{U_h^2}\right]^{-\frac{1}{2}}$$

ただし、

Uh: ガボールフィルタバンクの最大中心周波数
Ul: ガボールフィルタバンクの最小中心周波数

[0048] A texture feature vector is expressed by this several 6. Therefore, the degree HT of different kind of the texture of each field is computed by substituting the average of distance with the mean vector of all fields, and distance, and the standard deviation of distance for several 4.
[0049] Furthermore, the degree HS of different kind of the area of each field is computed by substituting the average of a difference with the average area of all fields, and a difference, and the standard deviation of a difference for several 4.
[0050] In the degree of different kind of the hole which the degree of different kind of the appearance of a field and a field include, it unifies and the degree HSh of different kind of the form of each field is obtained further again, as shown in several 8.
[0051]
[Equation 8]

$$HSh_i = \frac{SHD_i + HOD_i}{2}$$

ただし、

HSh_i: 領域iの形の異質度
SHD_i: 領域iの外形の異質度
HOD_i: 領域iの穴の異質度

[0052] Here, it is necessary to take into consideration both the difference in the configuration itself, and the difference in the way of being visible depended on rotating a figure about the difference in a form. Then, a configuration can be described using the P type Fourier descriptor which Yoshinori Kamisaka indicated to "the new Fourier descriptor applicable also to an open curve" in 1984/3 "a telecommunication society paper magazine." In this case, if the power to each frequency after the P type Fourier transform is detected, power is in agreement when the configuration of two figures is the same. Moreover, if a Fourier descriptor is used, it restricts to the time when a configuration and an angle of rotation are in agreement, and a Fourier descriptor is in agreement. Therefore, a Fourier descriptor and the power to each frequency can estimate how many configurations and appearance of two figures are in agreement.
[0053] As mentioned above, the feature of an appearance can be expressed by the vector which consists of a Fourier coefficient and power, and the difference of an appearance can be expressed using the Euclidean distance between these feature vectors. Therefore, the degree HSh of different kind of the appearance of each field is computed by substituting the average of distance with the mean vector of all fields, and distance, and the standard deviation of distance for several 4.
[0054] Next, the degree Hho of different kind of a hole is explained. In order to express the feature of the hole which a field includes, it is necessary to take into consideration not only the configuration of a hole but the number and hole site of a hole. Asking for the primary moment of a field can define the difference in a hole site. In addition, the primary moment is indicated by 1992 "an image-analysis handbook." Therefore, the feature of a hole was expressed by the vector as shown in several 9. In addition, turn of a hole is made into the order near the zero of a picture.
[0055]
[Equation 9]

$ho=[nh, M, ha_{11}, \cdots, ha_{1k}, hb_{21}, \cdots, hb_{2k}, \cdots, ha_{n1}, \cdots, ha_{nk}, hb_{n1}, \cdots, hb_{nk}]$

ただし、

ho : 穴の形の特徴ﾍﾞｸﾄﾙ
nh_i : 領域iの穴の数
M_i : 領域iのﾓｰﾒﾝﾄ
$ha_{i,1,j}$: 領域iの1番目の穴のj次のﾌｰﾘｴ係数
$hb_{i,1,j}$: 領域iの1番目の穴のj次のﾊﾟﾜｰ

[0056] Therefore, the degree Hho of different kind of the hole of each field is computed by substituting the average of distance with the mean vector of all fields, and distance, and the standard deviation of distance for several 4.
[0057] Moreover, the degree FP of [******] shown in several 2 can be defined as shown in several 10.
[0058]
[Equation 10]
$FP_{i,t}=wf_1 \cdot PC_i + wf_2 \cdot PT_i + wf_3 \cdot PS_{i,t}$

ただし、

$FP_{i,t}$: 領域iの時間tにおける特徴誘目度
PC_i : 領域iの色の誘目度
PT_i : 領域iの空間周波数の誘目度
$PS_{i,t}$: 領域iの面積の誘目度
wf_i : 重み係数

[0059] In several 10, the degree PC of a color of [**] is explained first.
[0060] the conventional research **** -- 3 lightness in which the one where 2 saturation in which 1 warm color is conspicuous from cold color is higher is conspicuous is reported for the higher one to be conspicuous 1) is reported for red not to be dependent on a background color and to be a color with the high degree of [**] by the experiment which **** and others conducted. Based on this result, in this example, it considers as the high hue whose red (0 R, G, B= 255, 0) is the degree of [**] most, and it is assumed that the degree of [**] is high, so that a hue is close to red. HSI indicated in detail by above-mentioned 1992 "an image-analysis handbook" in the hue here -- sugoroku -- when it asks using a pyramid color model, a red (0 R, G, B= 255, 0) hue is 0 That is, the point which a HSI congruence hexagon-head drill color model makes BK (black) a zero, and carries out a counter electrode to it is set to W (white). And the straight line which connects BK and W is a center mostly, and the hexagon which makes a vertex R (red), M (MAZENDA), B (blue), C (cyanogen), G (green), and Y (yellow) so that it may intersect perpendicularly with the straight line is formed. thus, formed HSI -- sugoroku -- the color space of a pyramid color model is used In addition, a color space means the space of the rectangular three-dimensions system of coordinates used in order to display a perceived color as one point of space. That is, in this example, the degree helium of a hue of [**] becomes large, so that a hue is close to 0, in order to search for a hue using a HSI congruence hexagon-head drill color model. For this reason, the degree helium of a hue of [**] is computed by the formula shown in several 11.
[0061]
[Equation 11]
$He = 1 - \frac{h}{\pi}$

ただし、

He : 色相誘目度
h : 領域の平均色相
(但し, $h>\pi$ の場合は$h=\pi-h$とする)

[0062] Moreover, about 2, an alignment-relation between saturation and the degree of [**] is. therefore -- this example -- the degree of saturation of [**] -- HSI -- sugoroku -- we decided to use the saturation in a pyramid color model itself

[0063] Furthermore, about 3, Semmelroth showed that the relation shown by several 12 was realized in the research on the relation between human being's amount of feelings, and lightness. In addition, the relation shown in several 12 is shown by Daisen, Imai, Wake and others in 1996 "a new edition feeling / consciousness psychology handbook."
[0064]

[Equation 12]

$$R = S^m + k|S - S_B|^n \qquad S \geqq S_B$$

$$R = S^m - k|S - S_B|^n \qquad S < S_B$$

ただし、

R : 感覚量
S : 視標輝度
S_B : 背景の輝度
k : 定数
m : 視標輝度への反応に対するﾍﾞｷ
n : 視標と背景の輝度差に対するﾍﾞｷ

[0065] Moreover, in several 12, suiting well with human being's amount of feelings in the case of k= 0.65, m= 0.4, and n= 0.2 was shown. Therefore, the formula shown by Semmelroth was made into the degree of lightness of [**] in this example.

[0066] Linear combination of the degree to three elements (a hue, saturation, lightness) of the above colors of [**] is carried out using several 13, and the degree PC of a color of [**] is defined.

[0067]

[Equation 13]

$$\frac{|He + S + I|}{3}$$

ただし、

He: 色相の誘目度
S : 彩度の誘目度
I : 明度の誘目度(数12のRと同値)

[0068] Next, the degree PT of a texture (spatial frequency) of [**] is explained.

[0069] In human being's visual system, it is shown clearly that it has the property of a band-pass machine in which sensitivity serves as the maximum by specific spatial frequency. Until now, Kubota, Nishizawa and others are formulizing visual spatial frequency characteristics like several 14. In addition, the formula of spatial frequency characteristics is indicated by Kubota and others in 1986/5 "a telecommunication society paper magazine" at "the 3-dimensional weighting-of-noise function of a television system, and application to the high quality TV." In addition, the unit of spatial frequency is changed into cpd (the number of cycles per 1 degree of visual senses) in several 14.

[0070]

[Equation 14]

$$V=(X,Y)=A\cdot\left[1-\frac{1}{1+\left(\frac{X+Y}{0.444}\right)^2}\right]/\left[1+\left(\frac{X+Y}{5\sqrt{d}}\right)^2\right]^2$$

$$A=\left[1+\left(\frac{f_m}{5\sqrt{d}}\right)^2\right]^2/\left[1-\frac{1}{1+\left(\frac{f_m}{0.444}\right)^2}\right]$$

ただし、

V : 視覚反応
X : 水平空間周波数
Y : 垂直空間周波数
A : 視覚反応のﾋﾟｰｸを1に規格化するための係数
d : 視距離(画面高の倍数)
f_m : ﾋﾟｰｸとなる空間周波数

[0071] The degree PT of a texture (spatial frequency) of [**] is defined by the visual-sense reaction V shown in this several 14.
[0072] Then, the degree PS of area of [**] is explained.
[0073] It is said that the direction of a near object is conspicuous from the center of a picture immediately after generally showing a picture, and if a point of regard is moved, the direction of the object near the point of regard which moved is said for the degree of [**] to become high. This is defined as the degree of a place of [**]. Thus, since the degree of each point of [**] models a bird clapper small gradually whenever it keeps away from the point focusing on a certain point (point of regard), the two-dimensional gauss function shown in several 15 is used. However, it is assumed that the center of a gauss function changes according to the time which the point of regard moved.
[0074]
[Equation 15]

$$E(x,y,\sigma,t)=\frac{1}{2\pi\sigma^2}\exp^{-\frac{(x-c_x(t))^2+(y-c_y(t))^2}{2\sigma^2}}$$

ただし、

x, y: 画素の座標
cx(t), cy(t): 中心座標
σ: 広がり係数

[0075] Here, it depends for the breadth coefficient sigma of a gauss function on the distance d from a view to a picture. That is, if Distance d becomes large, the range which can be observed at once will become large. Generally, human being's visible visual field is 20 - 30 degrees. Then, the breadth coefficient sigma when setting the screen quantity of the picture shown to H, and making a viewing distance into d-H can be defined like several 16. In addition, Pix is the vertical number of pixels and theta is 20 - 30 degrees (0.176<tan theta / 2< 0.268).
[0076]
[Equation 16]

$$\sigma=2\cdot Pix\cdot d\cdot\tan\frac{\theta}{2}$$

[0077] Here, if it takes into consideration that it is the set whose field is a pixel, the degree PS of the area of a field of [**] can be expressed by the sum of the degree of the place of the pixel of a field of [**]. Therefore, the degree PS of area of [**] is defined like several 17.
[0078]
[Equation 17]

$$PS_{i,t} = \sum_{x \in RX_i, y \in RY_i} E(x, y, \sigma, t)$$

ただし、

$PS_{i,t}$: 領域iの時間tにおける面積の誘目度
RX_i : 領域iの画素のX座標の集合
RY_i : 領域iの画素のY座標の集合

[0079] Thus, it asks for the drawing field where Step S37 estimates the degree of each drawing field of [**], and the degree of [**] serves as the maximum at Step S39 by the performance index of the defined degree of [**]. That is, a most conspicuous field is determined as an attention field. Therefore, an attention field as shown in drawing 4 (D) can be extracted.
[0080] In this example, in addition, each coefficient of the performance index of the degree of [**] wh1, wh2, wh3, [wh4] = [0.039, 0.010, 0.027, 0.020], and wf1, wf2 and = [0.132, 0.005, 0.100], and [wf3] [m, n] = [1.358, 4.250] are used. Several 14 theta was made into 20 degrees, and the viewing distance d was set to 1m, and the P type Fourier coefficient was carried out to to the 10th order.
[0081] Moreover, although the degree of [**] was evaluated using the eight above-mentioned physical features, this is for making two or more pictures which have all the features suit, and not necessarily needs to evaluate by this example about no physical features.
[0082] Then, at Step S41, it asks for the drawing field contiguous to the determined attention field, the color difference of an attention field and an adjoining-sheets field and the Euclidean distance of a texture feature vector are found, that whose Euclidean distance of a texture feature vector the color difference is less than 2.0, and is less than 0.3 is extracted with an attention field, and processing is ended. In addition, the color difference is CIE as are mentioned above and shows several 5. L* a* b* It asks by the color difference formula in a consciousness equal color space. Moreover, a texture feature vector is called for according to several 6.
[0083] It returns to drawing 1 and the attention field extracted with the attention field extractor 20 is given to composition cutoff equipment 22. Memory 24 is connected to composition cutoff equipment 22, and two or more data related with the photograph which the picture which the painter drew, and the photographer took in memory 24 are memorized. Composition cutoff equipment 22 cuts out an attention field from a subject-copy image with reference to the data memorized by memory 24 according to the composition of a painter or a photographer. Specifically, composition cutoff equipment 22 cuts off an attention field according to the flow view shown in drawing 7 . That is, if an attention field and an adjoining drawing field are extracted by the attention field extractor 20, composition cutoff equipment 22 starts processing, will be Step S51 and will ask for the periphery pixel of an attention field. That is, it asks for the edge of the extracted attention field. At continuing Step S53, by the technique of the paper which above-mentioned Kamisaka indicated, it asks for the P type Fourier coefficient to the 10th order, and an attention field is made into a configuration vector.
[0084] Here, the data memorized by memory 24 are the image data corresponding to criteria pictures, such as a pictures picture and a photograph, the data of the configuration vector which extracted the photographic subject from the pictures picture and photograph, and described a photographic subject's periphery configuration by the P type Fourier coefficient, and position data corresponding to the positional information of a photographic subject as shows drawing 8 . That is, a photographic subject's position data are data of the criteria composition of a criteria picture. The position of the zero (peak near the peak at the upper left of a picture) of b and a circumscription rectangle for the length of the side of a and length (w1, h1), [the length of the side beside the circumscription rectangle surrounding a photographic subject] It is expressed with a vector like several 18 when the position (position on the basis of the peak at the lower right of a picture) of the terminal point (peak near the peak at the lower right of a picture) of a circumscription rectangle is set to (w2, h2).
[0085]
[Equation 18] Position data [kx1, ky1, kx2, ky2] = [w1 /1 /2 / h2 [a and]/b] [a and h1] [b and w2]
Then, the youg lid distance of the configuration vector searched for at Step S53 by Step S55 and the configuration vector of a pictures picture or a photograph stored in memory 24 is found, that is, matching with the data of criteria composition is performed, and the data with which the configuration of a field was most similar at Step S57 are acquired. In other words, the position data of the photographic subject of a pictures picture with the smallest Euclidean distance are acquired. That is, the optimal image data for an attention field and an adjoining drawing field is chosen from two or more image data. And at Step S59, it asks for the circumscription rectangle of an attention field, and some subject-copy images are cut out from the zero and terminal point of a circumscription rectangle according to a-19

number. In addition, by several 19, * means multiplication.
[0086]
[Equation 19] Width of face of X1=x1-W*kx1Y1=y1-H*ky1X2=x2+W*kx2Y2=y2+W*ky2, however the circumscription rectangle of an attention field is set to W, and (x1, y1), and a terminal point are set [height] to (x2, y2) for the zero of H and a circumscription rectangle. Moreover, it considers as the height of a Y2= subject-copy image at the time of the width of face of X2= subject-copy image, and the height of a Y2> subject-copy image at the time of the width of face of Y1=1 and X2> subject-copy image at the time of X1=1 and Y1<1 at the time of X1<1. At this time, the peak coordinate of the rectangle (picture) to cut off is shown like several 20.
[0087]
[Equation 20] peak coordinate =[(X1, Y1) of the picture to cut off, (X1, Y2), and (X2, Y1) --] (X2, Y2)
Then, the result (picture) cut off at Step S61 is outputted, and processing is ended.
[0088] Therefore, according to the composition of a pictures picture as shown in drawing 9 (A), an attention field and an adjoining drawing field as shown in drawing 9 (B) can be cut off.
[0089] Since the field (attention field) which is most conspicuous from a subject-copy image according to the physical feature which was mentioned above is extracted according to this example, the attention field which suited human being's subjectivity can be extracted.
[0090] Moreover, a photograph since the extracted attention field was cut off according to the composition of the photograph which the pictures which the painter drew, and the photographer took, as if it photoed the photographic subject in the composition is generable. That is, it is automatic and the good composition of balance can be determined.
[0091] In addition, since this attention field extractor suits human being's subjectivity and can ask for the degree of a picture field of [**], it is applicable to equipment which performs objective evaluation of each field of the picture used as the candidate for evaluation, or an image which is conspicuous and performs weighting to the field according to whenever (the degree of [**]), for example in the quality evaluation of a digital compression picture and an image.
[0092] Moreover, in a printing field, it is applicable to the equipment which is automatic and can judge the field which can make light of problems, such as a printing gap, and the field which is not so according to the degree of [**] in automation of quality control. [0093] Furthermore, in the poster work used for a design field, especially an advertisement, it is applicable to the equipment by which it evaluates objective whether it is that the portion against which a company wants to appeal most is conspicuous.
[0094] Moreover, if it controls rotation of a camera, a tilt, and zoom using an attention field extractor in being the dress with which a user is conspicuous, a camera is automatic and it can run after a user. A user's snapshot can be taken, if it follows, for example, a shutter is turned off by predetermined timing. Furthermore, using the photoed picture (subject-copy image) to composition determination equipment, it is automatic and the snapshot of the good user of balance can be created by determining composition.

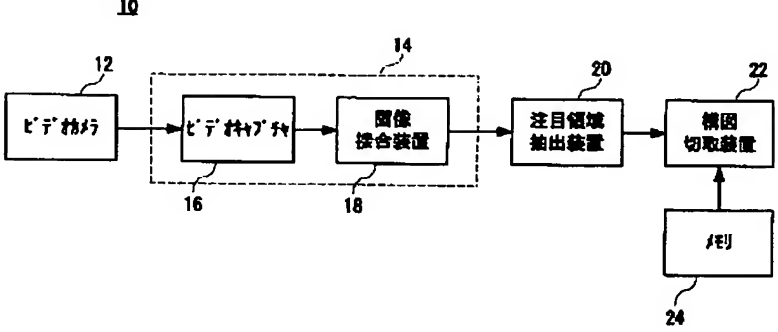
[Translation done.]

* NOTICES *

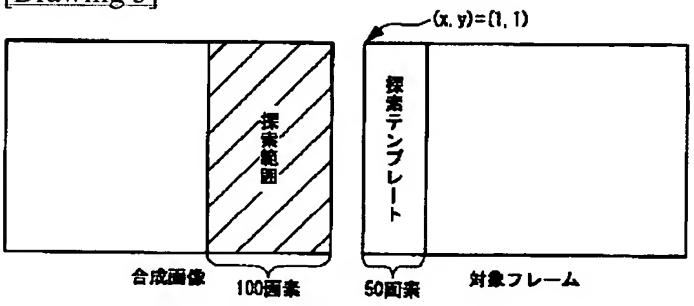
1.This document has been translated by computer. So the translation may not reflect the original precisely.
2.**** shows the word which can not be translated.
3.In the drawings, any words are not translated.

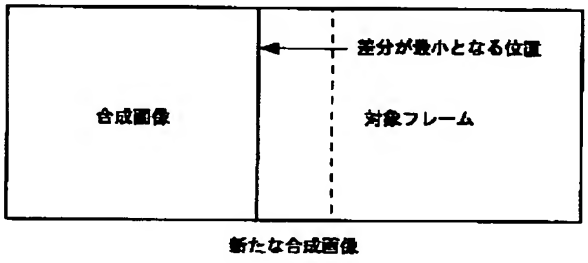
DRAWINGS

[Drawing 1]



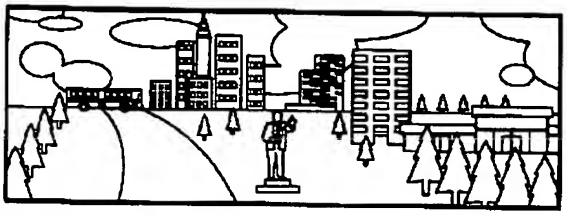
[Drawing 3]



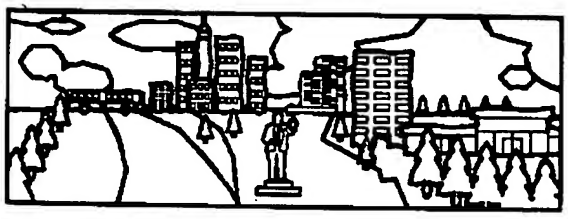


[Drawing 4]

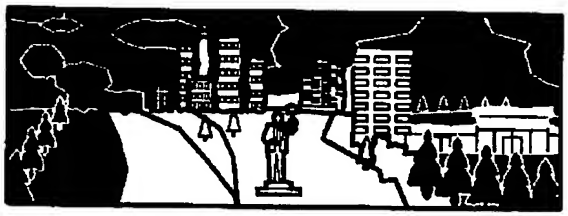
(A)ﾊﾟﾉﾗﾏ画像(原画像)



(B)領域分割結果

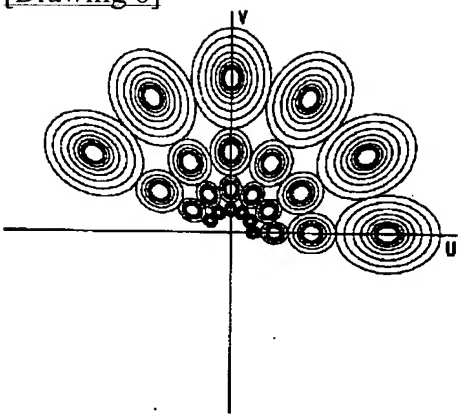


(C)囲領域抽出結果

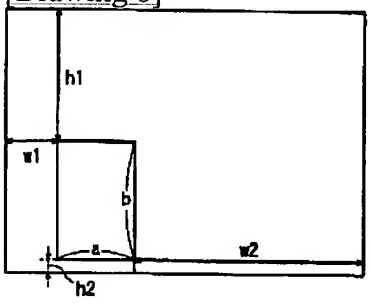


(D)注目領域抽出結果

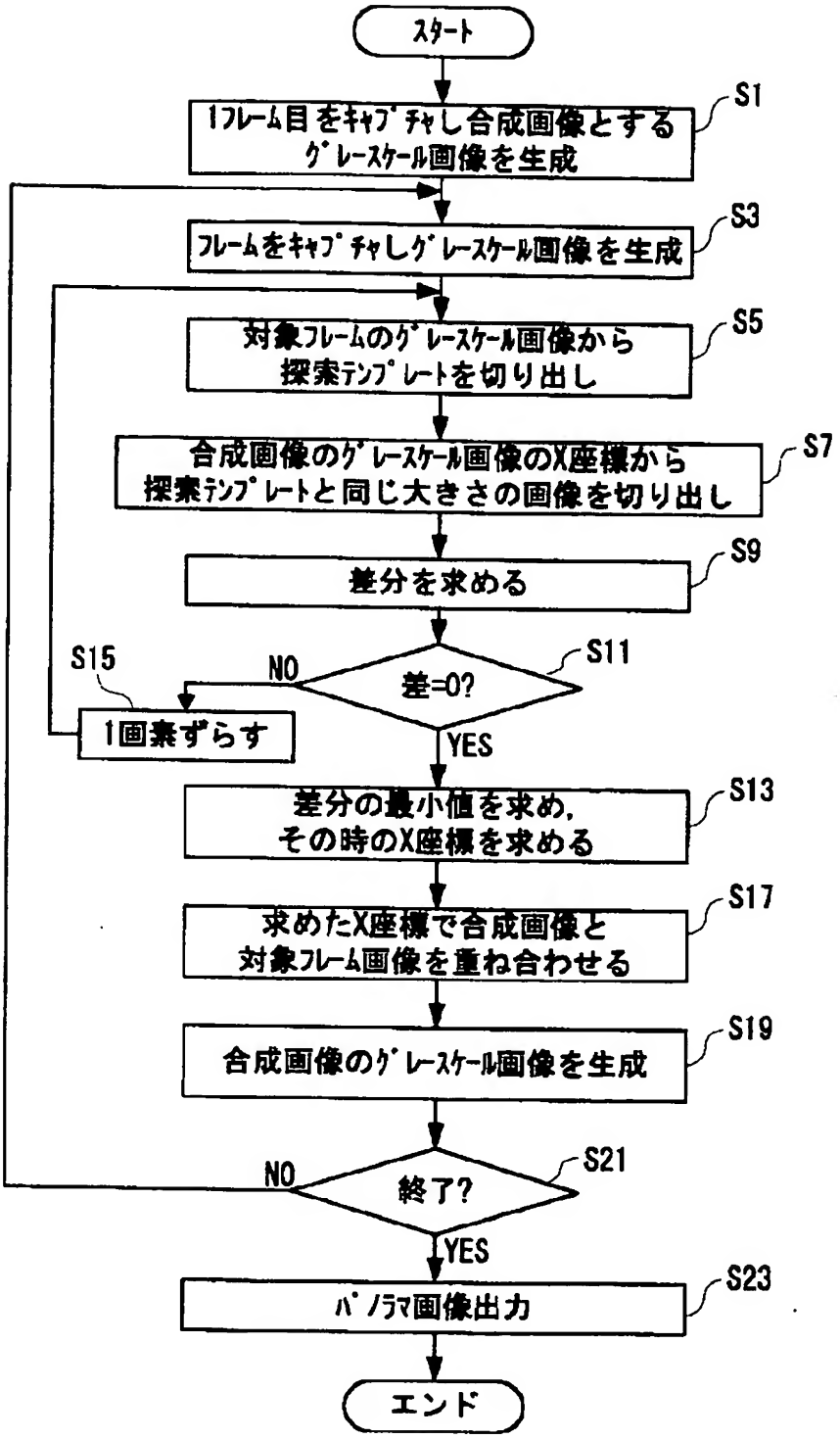
[Drawing 6]



[Drawing 8]

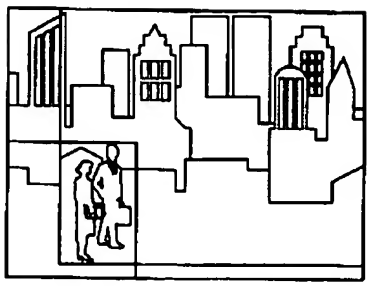


[Drawing 2]

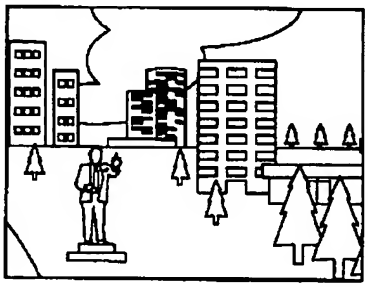

スタート
1フレーム目をキャプチャし合成画像とする
グレースケール画像を生成
S1
フレームをキャプチャしグレースケール画像を生成
S3
対象フレームのグレースケール画像から
探索テンプレートを切り出し
S5
合成画像のグレースケール画像のX座標から
探索テンプレートと同じ大きさの画像を切り出し
S7
差分を求める
S9
差=0?
S11
NO
YES
S15
1画素ずらす
差分の最小値を求め,
その時のX座標を求める
S13
求めたX座標で合成画像と
対象フレーム画像を重ね合わせる
S17
合成画像のグレースケール画像を生成
S19
終了?
S21
NO
YES
パノラマ画像出力
S23
エンド

[Drawing 9]

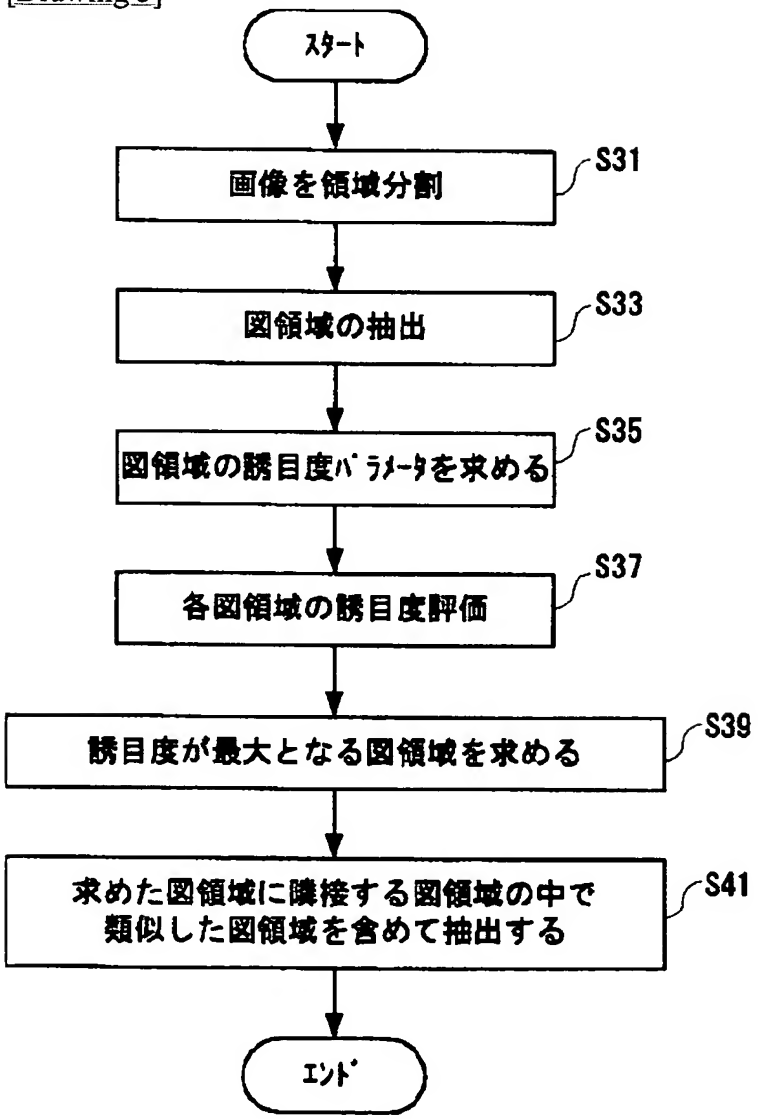
(A)取得した 図情報



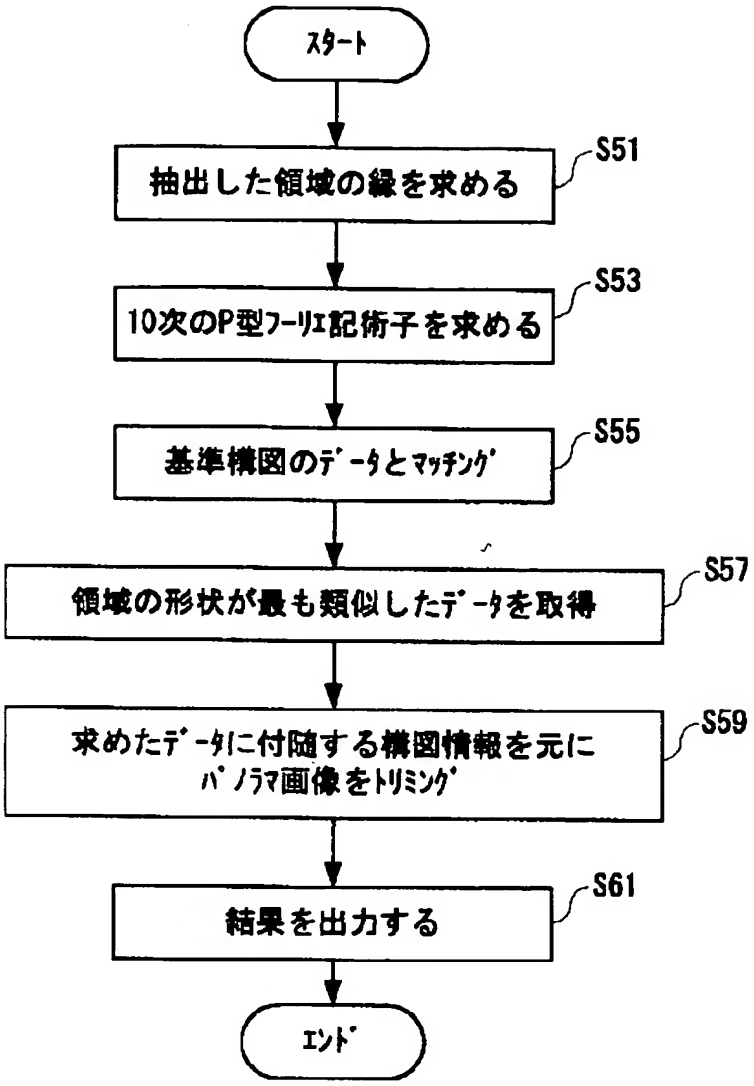
(B)切り取り結果



[Drawing 5]



[Drawing 7]



[Translation done.]

(19)日本国特許庁（ＪＰ）　(12)公開特許公報（Ａ）　(11)特許出願公開番号
特開2001－126070
（P2001－126070A）
(43)公開日　平成13年５月11日(2001.5.11)

(51)Int.Cl.⁷	識別記号	FI		テーマコード（参考）
Ｇ０６Ｔ　7/00		Ｈ０４Ｎ　1/387		５Ｂ０５７
1/00		5/232	Ｚ	５Ｃ０２２
Ｈ０４Ｎ　1/387		5/262		５Ｃ０２３
5/232		Ｇ０６Ｆ　15/70	３１０	５Ｃ０７６
5/262		15/66	４７０Ｊ	５Ｌ０９６

審査請求　有　請求項の数10　ＯＬ　（全 15 頁）　最終頁に続く

(21)出願番号　特願平11－306884

(22)出願日　平成11年10月28日(1999.10.28)

(71)出願人　595147700
株式会社エイ・ティ・アール知能映像通信研究所
京都府相楽郡精華町光台二丁目２番地２

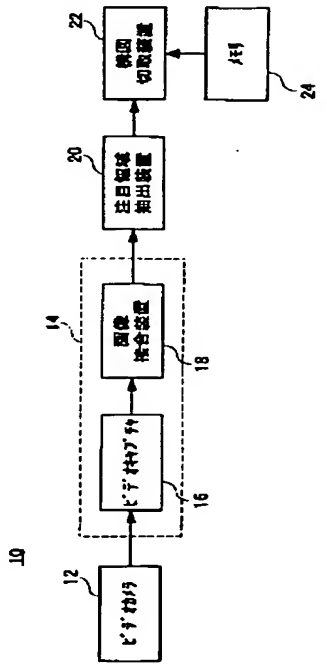
(72)発明者　田中　昭二
京都府相楽郡精華町大字乾谷小字三平谷５番地　株式会社エイ・ティ・アール知能映像通信研究所内

(74)代理人　100090181
弁理士　山田　義人

最終頁に続く

(54)【発明の名称】　注目領域抽出装置およびそれを用いた自動構図決定装置

(57)【要約】
【構成】　注目領域抽出装置およびそれを用いた自動構図決定装置１０は、画像生成装置１４を含み、画像生成装置１４はカメラ１２で撮影した映像からパノラマ画像の現画像を生成する。注目領域抽出装置２０は、画像生成装置１４から与えられる原画像から注目領域を抽出する。つまり、原画像の物理的特徴に従って人間の主観に合った評価をし、評価した結果に従って注目領域を抽出する。構図切り取り装置２２は、抽出した注目領域および隣接する画像領域を、メモリ２４に記憶された画家が描いた絵画や写真家が撮影した写真に関するデータを参照して、原画像から切り取る。つまり、絵画画像や写真画像と同じ構図で切り取ることができる。
【効果】　人間の主観に適合した注目領域を抽出することができる。また、バランスの良い構図を自動で決定することができる。



1

【特許請求の範囲】
【請求項１】原画像から注目領域を抜き出す注目領域抽出装置であって、
物理的特徴に基づいて誘目度を評価する評価手段、および前記評価手段の評価結果に応じて前記注目領域を抽出する抽出手段を備える、注目領域抽出装置。
【請求項２】前記物理的特徴は色の異質度を含む、請求項１記載の注目領域抽出装置。
【請求項３】前記物理的特徴は形の異質度、面積の異質度およびテクスチャの異質度をさらに含み、
前記評価手段は、４つの異質度のうち少なくとも１つの異質度に基づいて前記誘目度を評価する、請求項２記載の注目領域抽出装置。
【請求項４】前記物理的特徴は色をさらに含む、請求項１ないし３のいずれかに記載の注目領域抽出装置。
【請求項５】前記物理的特徴は前記原画像における領域の面積および空間周波数をさらに含む、請求項４記載の注目領域抽出装置。
【請求項６】所望の映像を撮影する撮影手段、および前記映像に基づいて前記原画像を生成する画像生成手段をさらに備える、請求項１ないし５のいずれかに記載の注目領域抽出装置。
【請求項７】前記撮影手段はカメラ、および前記カメラを回転駆動する駆動手段を含み、
前記画像生成手段は前記カメラで撮影される前記映像を１フレーム毎に接合する接合手段を含む、請求項６記載の注目領域抽出装置。
【請求項８】請求項１ないし７のいずれかに記載の注目領域抽出装置を用いた自動構図決定装置であって、
基準構図を有する基準画像に関するデータを保持する保持手段、および前記基準構図を参照して注目領域の画像を原画像から切り取る切取手段を備える、自動構図決定装置。
【請求項９】前記保持手段は複数の前記基準画像に関するデータを保持し、
前記注目領域の画像に適合した前記データを選択する選択手段をさらに備える、請求項８記載の自動構図決定装置。
【請求項１０】前記データは、少なくとも前記基準画像に対応する画像データ、被写体の形状データおよび被写体の位置データを含む、請求項８または９記載の自動構図決定装置。
【発明の詳細な説明】
【０００１】
【産業上の利用分野】この発明は注目領域抽出装置およびそれを用いた自動構図決定装置に関し、特に例えば原画像から注目領域を抽出し、抽出した注目領域を画家や写真家などの専門家が制作した絵の構図に適合して収まるように原画像から切り取る、注目領域抽出装置およびそれを用いた自動構図決定装置に関する。

2

【０００２】ここで、注目領域とは、画像や映像において人が注目する領域をいう。
【０００３】
【従来の技術】この種の従来技術として、画像から観測者が注目領域を抽出する手法に関しては、種々の方法が提案されている。（１）Milanese,Itti およびKochらは、画像中の不連続部分を注目領域と仮定し、画像から得られる複数の物理的特徴に対応する特徴マップ（濃淡画像）を生成し、次に、各特徴マップの不連続部分を求め、それらを統合したものを注目領域として抽出している。
【０００４】（２）Milaneseらは、各特徴マップの不連続部分を大きさが固定された複数のDifference-of-ofiented-Gaussiansフィルタを用いて特徴マップをフィルタリングし、出力が最大となるフィルタリング結果を選択して求めた。Itti, Kochらは各特徴マップを各々の平均値との二乗誤差で正規化し、全ての特徴マップを線形結合により統合した。そして、Difference of Gaussianフィルタにより統合特徴マップを再帰的にフィルタリングし、最終的に得られるフィルタリング結果の局所的なピークを注目領域として抽出していた。
【０００５】このように、MilaneseやIttiらは、フィルタリングや弛緩法などの画素レベルの処理により注目領域を抽出していた。
【０００６】（３）Martinや竹内らは、画像から得られる輝度情報をシャノンの情報理論に基づいて評価し、その結果得られる情報量の高い部分を注目領域としていた。この方式では輝度値の分散が大きい領域つまり複雑に見える領域や明るい領域が主に抽出される。
【０００７】また、従来のカメラでは、人間が被写体および被写体の構図を手動で決定していた。
【０００８】
【発明が解決しようとする課題】しかし、（１）や（２）の手法では、画像によって注目領域の大きさが異なるため、大きさが固定されたフィルタを用いて注目領域を的確に抽出するのが困難であった。また、（３）の手法では、たとえば複雑な模様の壁の前に置かれた黒色の花瓶の絵のように、複雑に見える領域、あるいは明るい領域が必ずしも注目領域と一致するとは言えない絵の場合には、注目領域を的確に抽出するのが困難であった。さらに、このような提案では、観測者（人間）の主観と抽出結果との照合を行った例が少なく、実際に人間の主観に適合する注目領域を抽出できるかどうかが疑問であった。
【０００９】また、従来のカメラでは写真に対するセンスのない一般の人が構図を決定した場合、必ずしもバランスの良い写真を撮影できるとは限らない。
【００１０】それゆえに、この発明の主たる目的は、人間の主観に適合した注目領域を的確に抽出することができる、注目領域抽出装置を提供することである。

【００１１】また、この発明の他の目的は、バランスの良い構図を自動で決定することができる、自動構図決定装置を提供することである。

【００１２】

【課題を解決するための手段】第１の発明は、原画像から注目領域を抜き出す注目領域抽出装置であって、物理的特徴に基づいて誘目度（目立ち度）を評価する評価手段、および評価手段の評価結果に応じて注目領域を抽出する抽出手段を備える、注目領域抽出装置である。

【００１３】第２の発明は、第１の発明に記載の注目領域抽出装置を用いた自動構図決定装置であって、基準構図を有する基準画像に関するデータを保持する保持手段、および基準構図を参照して注目領域の画像を原画像から切り取る切取手段を備える、自動構図決定装置である。

【００１４】

【作用】第１の発明の注目領域抽出装置では、評価手段が原画像の物理的特徴に従って誘目度を評価する。ここで、誘目度とは、人間の主観に合ったパラメータをいう。抽出手段は、評価結果から一番目立つ領域を注目領域として抽出する。つまり、評価手段は物理的特徴に従って人間の主観に合った評価をするので、人間の主観に適合した注目領域を抽出することができる。

【００１５】たとえば、物理的特徴が色の異質度を含む場合、各領域の色の違いに基づいて誘目度を評価することができる。

【００１６】また、物理的特徴が、色の異質度に加えて、形の異質度、面積の異質度およびテクスチャ（模様）の異質度をさらに含むので、この４つの異質度の少なくとも１つの異質度に基づいて誘目度を評価すれば、原画像の特徴に応じて的確に誘目度を評価することができる。

【００１７】また、色の３要素（色相、彩度、明度）についても評価する場合であれば、人間の主観による目立つ色（赤色）に近い領域を最も目立つ領域と評価することができる。

【００１８】さらに、空間周波数や原画像における各領域の面積についても評価すれば、最も目立つ領域の評価をさらに的確に判断することができる。

【００１９】また、撮影手段で所望の映像を撮影し、たとえば撮影した映像に対応する画像を合成し、原画像を生成することもできる。

【００２０】たとえば、撮影手段に含まれるカメラの位置および高さを固定し、３６０°回転して撮影できるようにしておけば、接合手段が１フレーム毎の画像を接合することにより、３６０°の範囲内でパノラマ画像の原画像を生成することができる。

【００２１】第２の発明の自動構図決定装置では、保持手段が、たとえば画家が描いた絵画や写真家が撮影した写真に対応する絵画画像および写真画像を基準画像とし、この基準画像に関するデータを保持する。切取手段は、この基準画像に関するデータを参照して注目領域の画像を原画像から切り取るので、バランスの良い構図を自動で決定することができる。

【００２２】また、複数の基準画像に関するデータを保持しておけば、選択手段が注目領域の画像に適した構図を選択できるので、あらゆる画像についてバランスの良い構図を決定することができる。

【００２３】上述のような基準画像に関するデータは、少なくとも基準画像に対応する画像データ、被写体の形状データおよび被写体の位置データを含むので、注目領域に適した構図を選択し、バランスの良い構図を自動で決定できる。つまり、あたかも被写体をその構図で撮影したかのような写真を生成することができる。

【００２４】

【発明の効果】この発明によれば、人間の主観に合った物理的特徴の評価結果に従って注目領域を抽出するので、人間の主観に適合した注目領域を抽出することができる。

【００２５】他の発明によれば、基準画像の基準構図を用いて注目領域を原画像から切り取るので、バランスの良い構図を自動で決定することができる。

【００２６】この発明の上述の目的，その他の目的，特徴および利点は、図面を参照して行う以下の実施例の詳細な説明から一層明らかとなろう。

【００２７】

【実施例】図１を参照して、この実施例の自動構図決定装置１０は、ビデオカメラ（以下、単に「カメラ」という。）１２を含む。カメラ１２は、たとえば三脚などを用いて撮影位置（位置および高さ）が固定され、３６０°回転して撮影できる。カメラ１２で撮影した映像が画像生成装置１４に入力され、撮影した映像から原画像としてのパノラマ画像が生成される。なお、カメラ１２は３６０°回転できるので、画像生成装置１４は３６０°の範囲内でパノラマ画像を生成できる。

【００２８】画像生成装置１４は、ビデオキャプチャ１６および画像接合装置１８を含み、図２に示すフロー図に従ってパノラマ画像を生成する。つまり、カメラ１２で撮影が開始されると、画像生成装置１４は処理を開始し、ステップＳ１で、ビデオキャプチャ１６を用いて最初のビデオフレーム（１フレーム目）をキャプチャし、その１フレーム目に対応する画像から合成画像を生成した後、画像接合装置１８を用いてその合成画像からグレースケール画像を生成する。

【００２９】続くステップＳ３では、次のフレームをキャプチャし、１フレーム目と同様にグレースケール画像を生成する。ステップＳ５では、現フレーム（対象フレーム）に対応する合成画像をどの位置から合成すればよいかを決定するため、図３に示すような対象フレームのグレースケール画像から探索テンプレートを切り出す。

この探索テンプレートの幅は５０画素であり、その高さはフレーム画像と同じである。また、この探索プレートは、対象フレームの原点から切り出される。

【００３０】続いてステップＳ７では、合成画像のグレースケール画像から探索範囲を設定し、探索範囲から探索テンプレートと同じ大きさの画像を切り出す。つまり、探索範囲の幅は１００画素であり、その高さはフレーム画像と同じである。また、フレーム画像の左上を原点（（ｘ，ｙ）＝（１，１））としたとき、探索プレートのｘ座標はフレーム画像の幅から１００だけ引いた位置であり、そのｙ座標は１である。

【００３１】次に、ステップＳ９では、探索範囲から探索テンプレートと同じ大きさの画像を切り出し、切り出した画像および探索プレートに対応する互いの画素値の差の絶対値を算出する。続くステップＳ１１では、画素値の差が最小（０）であるかどうかを判断する。ステップＳ１１で“ＹＥＳ”であれば、切り出した画像が探索プレートと同じ大きさであると判断し、ステップＳ１３に進む。一方、ステップＳ１１で“ＮＯ”であれば、切り出した画像と探索プレートとの大きさが異なると判断し、ステップＳ１５で探索範囲をｘ軸の正の方向に１画素ずらしてから、ステップＳ５に戻る。このように、１画素ずつずらしながら、切り出した画像と探索テンプレートとが同じ大きさになるまで、処理が繰り返される。

【００３２】ステップＳ１３では、差の絶対値が最小（０）である時のＸ座標を算出する。続くステップＳ１７では、算出したＸ座標で合成画像と対象フレーム画像とを重ね合わせ、新たな合成画像を生成する。そして、ステップＳ１９では生成した合成画像からグレースケール画像を生成し、ステップＳ２１でパノラマ画像が生成されたかどうかを判断する。ステップＳ２１で“ＮＯ”であれば、つまりパノラマ画像が生成されてなければ、ステップＳ３に戻る。一方、ステップＳ２１で“ＹＥＳ”であれば、図４（Ａ）に示すようなパノラマ画像（原画像）が生成されたと判断し、ステップＳ２３で原画像を後段の注目領域抽出装置２０（図１）に出力し、＊処理を終了する。

【００３３】図１に戻って、注目領域抽出装置２０は、与えられた原画像から最も目立つ領域（注目領域）を抽出する。具体的には、注目領域抽出装置２０は、図５に示すフロー図に従って処理をする。つまり、原画像が画像生成装置１４から与えられると、注目領域抽出装置２０は処理を開始し、ステップＳ３１で原画像を領域分割する。具体的には、図４（Ｂ）に示すように、原画像を図領域と絵領域に分割する。この領域分割の方法には、１９９７ＩＥＥＥにおいてW.Y.MaやB.S.Manjunath らが「Edge Flow:A Framework of Boundary Detection and Image Segmentation」に記載した“edge flow”に基づく境界検出方法が適用される。簡単に説明すると、この方法は画像の各場所における色や模様の変化方向を求め、変化方向と変化の強さ（大小）からなるedge flowベクトルを決定する。そして、反復処理によりedge flowベクトルを各ベクトルの方向へ伝搬し、最終的ベクトル同士がぶつかり合う場所を各領域の境界線とする。

【００３４】したがって、ステップＳ３３では、図４（Ｃ）に示すような分割した図領域を抽出し、ステップＳ３５で図領域の誘目度を評価する。つまり、図領域の誘目度パラメータを求める。ここで、発明者が行った主観評価実験により、誘目度の評価に必要な物理的特徴は、色の異質度，テクスチャの異質度，形の異質度および面積の異質度，色，空間周波数および面積であることが分かった。また、一般に人間の主観評価結果と物理的特性との関係はＳ字曲線で表されることが多いため、誘目度の評価には数１に示すベータ関数が用いられる。

【００３５】

【数１】

$$H(m,n)=\int_0^1 h^{m-1}(1-h^{n-1})\,dh$$

【００３６】このベータ関数を用いて、数２に示すような誘目度の評価関数が定義される。

【００３７】

【数２】

$$A_{i,t}=\int_0^1 a_{i,t}^{m-1}(1-a_{i,t}^{n-1})\,da_{i,t}$$

$$a_i=HET_i+FP_{i,t}$$

ただし、

$A_{i,t}$：　領域iの時間tにおける誘目度
$a_{i,t}$：　領域iの時間tにおける特徴量統合値
HET_i：　領域iの異質性誘目度
$FP_{i,t}$：　領域iの時間tにおける特徴誘目度
i：　領域i
t：　時間
m，n：　立ち上がり，飽和パラメータ

【００３８】また、数２に示す各領域の上記４つの異質度に対する誘目度ＨＥＴは数３で定義される。

【００３９】

【数３】

$$HET_i = wh_1 \cdot HC_i + wh_2 \cdot HT_i + wh_3 \cdot HS_i + wh_4 \cdot HSh_i$$

ただし、

HET_i ： 領域iの異質性誘目度
HC_i ： 領域iの色の異質度
HT_i ： 領域iのﾃｸｽﾁｬの異質度
HS_i ： 領域iの面積の異質度
HSh_i ： 領域iの形の異質度
wh_i ： 重み係数

【００４０】さらに、各物理的特徴の異質度Ｈは、特徴値と全領域の平均特徴値との差をｄ，差ｄの平均値をｄｍ，差ｄの標準偏差をｓｔｄとしたとき、数４に従って算出される。

【００４１】

【数４】

$$H(d, dm, std) = \left| \frac{d-dm}{std} \right|$$

【００４２】つまり、各領域の色の異質度ＨＣは、ＣＩＥ Ｌ＊ａ＊ｂ＊知覚均等色空間における色差式を用いて領域の平均色と全領域の平均色との色差、色差平均、色差の標準偏差を算出し、算出した結果を数４に代入して求められる。なお、色差式は、１９９４「色彩科学ハンドブック」に詳細に記載されている。具体的には、画像の各画素の色情報（Ｒ，Ｇ，Ｂ）を均等知覚色空間Ｌ＊ａ＊ｂ＊に変換し、その空間におけるユークリッド距離を色差としている。つまり、色差式は、数５のように示される。

【００４３】

【数５】

$$\Delta Eab^* = [(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)]^{\frac{1}{2}}$$

【００４４】次に、テクスチャの異質度ＨＴについて説明する。テクスチャは、１９９６ＩＥＥＥにおいてB.S. Manjunath や W.Y.Ma らが「Texture Features for Browsing and Retrieval of Image Data」に記載したテクスチャ特徴ベクトルで表現され、またテクスチャ間の差はテクスチャ特徴ベクトル間のユークリッド距離によって表現される。

【００４５】ここで、テクスチャ特徴ベクトルは、大きさと方向の異なる複数のガボールフィルタで構成したガボールフィルタバンクにより画像をフィルタリングしたときの応答を要素とするベクトルで表現される。ただし、複数のガボールフィルタによりフィルタリングした場合は、各応答には直交性がないため、フィルタリングした結果には冗長な情報が含まれる可能性がある。そこで、ガボールフィルタバンク内の各フィルタのパラメータは、上述の１９９６ＩＥＥＥに記載されている手法によって決定される。具体的には、数６で示されるガボールフィルタのθ、ａ、σuおよびσvを数７で求める。なお、この手法は、図６に示すように、隣接するフィルタがHalf-Peak で接するようにフィルタのスケール（大きさ）および方位パラメータを定めるものであり、テクスチャ特徴を表現するために、２４個のフィルタ（４スケール、６方向）が用いられる。

【００４６】

【数６】

9 10

$$G_{mn}(x,y)=a^{-m}\cdot G(x',y')$$

$$G(x,y)=\left(\frac{1}{2\pi\sigma_x\sigma_y}\right)\exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)\right]\cdot\exp(2\pi jUx)$$

$$U=a^{m}$$

$$x'=x\cos\theta+y\sin\theta$$

$$y'=-x\sin\theta+y\cos\theta$$

$$\sigma_x=\frac{1}{2\pi\sigma_u}$$

$$\sigma_y=\frac{1}{2\pi\sigma_v}$$

ただし、

Gmn(x, y)： スケールID=m, 方位ID=nの時のガボールフィルタ
m ： 0 .. S-1
n ： 0 .. K-1
S ： スケール数
K ： 方位数

【００４７】
【数７】

$$\theta=\frac{n\pi}{K}$$

$$a=(U_h/U_l)^{\frac{1}{S-1}}$$

$$\sigma_u=\frac{(a-1)U_h}{(a+1)\sqrt{2\ln 2}}$$

$$\sigma_v=\tan\left(\frac{\pi}{2k}\right)\left[U_h-2\ln 2\left(\frac{\sigma_u^2}{U_h}\right)\right]\left[2\ln 2-\frac{(2\ln 2)^2\sigma_u^2}{U_h^2}\right]^{\frac{1}{2}}$$

ただし、

Uh： ガボールフィルタバンクの最大中心周波数
Ul： ガボールフィルタバンクの最小中心周波数

【００４８】この数６によって、テクスチャ特徴ベクトルが表現される。したがって、各領域のテクスチャの異質度ＨＴは、全領域の平均ベクトルとの距離、距離の平均値、距離の標準偏差を数４に代入することにより算出される。

【００４９】さらに、各領域の面積の異質度ＨＳは、全領域の平均面積との差、差の平均値、差の標準偏差を数４に代入することにより算出される。

【００５０】さらにまた、各領域の形の異質度ＨＳｈは、領域の外形の異質度と、領域が包含する穴の異質度を数８に示すように統合して得られる。

【００５１】
【数８】

$$HSh_i=\frac{SHD_i+HOD_i}{2}$$

ただし、

HSh_i： 領域iの形の異質度
SHD_i： 領域iの外形の異質度
HOD_i： 領域iの穴の異質度

【００５２】ここで、形の違いについては、形状そのものの違いと、図形を回転することによる見え方の違いとの両方を考慮する必要がある。そこで、１９８４／３「電気通信学会論文誌」において上坂吉則が「開曲線にも適用できる新しいフーリエ記述子」に記載したＰ型フーリエ記述子を用いて、形状を記述することができる。この場合、Ｐ型フーリエ変換後の各周波数に対するパワーを検出すれば、２つの図形の形状が同じときには、パワーが一致する。また、フーリエ記述子を用いれば、形状と回転角が一致するときに限り、フーリエ記述子が一

11

致する。したがって、フーリエ記述子と各周波数に対するパワーとによって、２つの図形の形状と外見とがどの程度一致するかを評価することができる。

【００５３】以上から、外形の特徴はフーリエ係数とパワーとからなるベクトルで表現でき、また外形の差は同特徴ベクトル間のユークリッド距離を用いて表現できる。したがって、各領域の外形の異質度ＨＳｈは、全領域の平均ベクトルとの距離、距離の平均値、距離の標準偏差を数４に代入して算出される。

【００５４】次に、穴の異質度Ｈｈｏに関して説明す

12

＊る。領域が包含する穴の特徴を表現するためには、穴の形状だけでなく、穴の数および穴の位置も考慮する必要がある。穴の位置の違いは、領域の一次モーメントを求めることで定義できる。なお、一次モーメントは、１９９２「画像解析ハンドブック」に記載されている。したがって、穴の特徴は、数９に示すようなベクトルで表現した。なお、穴の順番は、画像の原点に近い順にしてある。

【００５５】

【数９】

$$ho=[nh, M, ha_{11}, \cdots, ha_{1k}, hb_{21}, \cdots, hb_{2k_2}, \cdots, ha_{n1}, \cdots, ha_{nk}, hb_{n1}, \cdots, hb_{nk}]$$

ただし、

ho ：　穴の形の特徴ベクトル
nh_i ：　領域ｉの穴の数
M_i ：　領域ｉのモーメント
$ha_{i,1,j}$ ：　領域ｉの1番目の穴のｊ次のフーリエ係数
$hb_{i,1,j}$ ：　領域ｉの1番目の穴のｊ次のパワー

【００５６】したがって、各領域の穴の異質度Ｈｈｏは、全領域の平均ベクトルとの距離、距離の平均値、距離の標準偏差を数４に代入して算出される。

【００５７】また、数２に示す特徴誘目度ＦＰは、数１０に示すように定義できる。

【００５８】

【数１０】

$$FP_{i,t}=wf_1\cdot PC_i+wf_2\cdot PT_i+wf_3\cdot PS_{i,t}$$

ただし、

$FP_{i,t}$ ：　領域ｉの時間tにおける特徴誘目度
PC_i ：　領域ｉの色の誘目度
PT_i ：　領域ｉの空間周波数の誘目度
$PS_{i,t}$ ：　領域ｉの面積の誘目度
wf_i ：　重み係数

【００５９】数１０において、まず色の誘目度ＰＣについて説明する。

【００６０】従来研究おいて、１）暖色の方が寒色よりも目立つ、２）彩度が高いほうが目立つ、３）明度が高いほうが目立つと報告されている。１）に関しては、神作らが行った実験によって、赤色は背景色に依存せず誘目度の高い色であると報告されている。この結果に基づき、この実施例では、赤色（Ｒ，Ｇ，Ｂ＝２５５，０，０）が最も誘目度の高い色相とし、色相が赤色に近いほど誘目度が高いと仮定する。ここで、色相を上述の１９９２「画像解析ハンドブック」に詳細に記載されたＨＳＩ双六角錐カラーモデルを用いて求めると、赤色（Ｒ，Ｇ，Ｂ＝２５５，０，０）の色相は０である。つまり、ＨＳＩ双六角錐カラーモデルは、ＢＫ（黒）を原点とし、それに対極する点がＷ（白）とされる。そして、ＢＫとＷとを結ぶ直線のほぼ中央であり、かつその直線に直交するようにＲ（赤），Ｍ（マゼンダ），Ｂ（青），Ｃ（シアン），Ｇ（緑），Ｙ（イエロ）を頂点とする六角形が形成される。このように形成されたＨＳＩ双六角錐カラーモデルの色空間が用いられる。なお、色空間とは、知覚色を空間の一点として表示するために用いる直交三次元座標系の空間をいう。つまり、この実施例ではＨＳＩ双六角錐カラーモデルを用いて色相を求めるため、色相が０に近いほど色相の誘目度Ｈｅが大きくなる。このため、色相の誘目度Ｈｅは、数１１に示す式によって算出される。

【００６１】

【数１１】

$$He=1-\frac{h}{\pi}$$

ただし、

He ：　色相誘目度
h ：　領域の平均色相
（但し，h>πの場合はh=π-hとする）

【００６２】また、２）に関しては、彩度と誘目度の間には線形的な関係がある。したがって、この実施例では、彩度の誘目度は、ＨＳＩ双六角錐カラーモデルにおける彩度そのものを用いることとした。

【００６３】さらに、３）に関しては、人間の感覚量と明度との関係に関する研究において、Semmelrothは数１２で示す関係が成り立つことを示した。なお、数１２に示す関係は、１９９６「新編　感覚・知覚　心理学ハンドブック」において、大山，今井，和気らによって示される。

【００６４】

【数１２】

$$R = S^m + k|S - S_B|^n \qquad S \geqq S_B$$

$$R = S^m - k|S - S_B|^n \qquad S < S_B$$

ただし、

R：　感覚量
S：　視標輝度
S_B：　背景の輝度
k：　定数
m：　視標輝度への反応に対するﾍﾞｷ
n：　視標と背景の輝度差に対するﾍﾞｷ

【００６５】また、数１２において、ｋ＝０．６５，ｍ＝０．４，ｎ＝０．２の場合に、人間の感覚量とよく合うことが示された。したがって、この実施例では、Semmelrothによって示された式を明度の誘目度とした。

【００６６】上述のような色の３要素（色相、彩度、明度）に対する誘目度を数１３を用いて線形結合し、色の誘目度ＰＣが定義される。

【００６７】

【数１３】

$$\frac{He+S+I}{3}$$

ただし、

He：　色相の誘目度
S：　彩度の誘目度
I：　明度の誘目度（数12のRと同値）

＊【００６８】次に、テクスチャ（空間周波数）の誘目度ＰＴについて説明する。

【００６９】人間の視覚系においては、特定の空間周波数で感度が最大となるような帯域通過器の性質を持つことが明らかにされている。これまで、久保田や西澤らは、視覚の空間周波数特性を数１４のように定式化している。なお、空間周波数特性の式は、１９８６／５「電気通信学会論文誌」において、久保田らによって「テレビジョン系の３次元雑音評価関数とその高品位テレビへの応用」に記載されている。なお、数１４では、空間周波数の単位をｃｐｄ（視覚１° あたりのサイクル数）に変換してある。

【００７０】

【数１４】

$$V=(X,Y)=A\cdot\left[1-\frac{1}{1+\left(\frac{X+Y}{0.444}\right)^2}\right]\Big/\left[1+\left(\frac{X+Y}{5\sqrt{d}}\right)^2\right]^2$$

$$A=\left[1+\left(\frac{f_m}{5\sqrt{d}}\right)^2\right]^2\Big/\left[1-\frac{1}{1+\left(\frac{f_m}{0.444}\right)^2}\right]$$

ただし、

V：　視覚反応
X：　水平空間周波数
Y：　垂直空間周波数
A：　視覚反応のﾋﾟｰｸを1に規格化するための係数
d：　視距離（画面高の倍数）
f_m：　ﾋﾟｰｸとなる空間周波数

【００７１】この数１４に示す視覚反応Ｖによって、テクスチャ（空間周波数）の誘目度ＰＴが定義される。

【００７２】続いて、面積の誘目度ＰＳについて説明する。

【００７３】一般に画像が提示された直後は、画像の中心から近いオブジェクトの方が目立つと言われており、また、注視点が移動されると、移動した注視点に近いオブジェクトの方が誘目度が高くなると言われている。これを、場の誘目度と定義する。このように、ある点（注視点）を中心とし、その点から遠ざかる毎に各点の誘目度が徐々に小さくなることをモデル化するために、数１５に示す２次元ガウス関数が用いられる。ただし、注視

点が移動した時間に応じてガウス関数の中心が変化すると仮定してある。

【００７４】

【数１５】

$$E(x, y, \sigma, t) = \frac{1}{2\pi\sigma^2} \exp^{-\frac{(x-c_x(t))^2+(y-c_y(t))^2}{2\sigma^2}}$$

ただし、

x, y:	画素の座標
cx(t), cy(t):	中心座標
σ:	広がり係数

【００７５】ここで、ガウス関数の広がり係数σは視点から画像までの距離ｄに依存する。つまり、距離ｄが大きくなれば、一度に観測可能な範囲が広くなる。一般に、人間の可視視野は２０度～３０度である。そこで、提示される画像の画面高をＨとし、視距離をｄ・Ｈとしたときの広がり係数σを数１６のように定義することができる。なお、Ｐｉｘは垂直方向の画素数であり、θは２０度～３０度（０．１７６＜tan θ／２＜０．２６８）である。

【００７６】

【数１６】

$$\sigma = 2 \cdot Pix \cdot d \cdot \tan\frac{\theta}{2}$$

【００７７】ここで、領域が画素の集合であることを考慮すれば、領域の面積の誘目度ＰＳは、領域の画素の場の誘目度の和で表現可能である。したがって、面積の誘目度ＰＳは数１７のように定義される。

【００７８】

【数１７】

$$PS_{i,t} = \sum_{x \in RX_i, y \in RY_i} E(x, y, \sigma, t)$$

ただし、

$PS_{i,t}$:	領域ｉの時間ｔにおける面積の誘目度
RX_i:	領域ｉの画素のＸ座標の集合
RY_i:	領域ｉの画素のＹ座標の集合

【００７９】このように、定義した誘目度の評価関数により、ステップＳ３７で各図領域の誘目度を評価し、ステップＳ３９で誘目度が最大となる図領域を求める。つまり、最も目立つ領域を注目領域に決定する。したがって、図４（Ｄ）に示すような注目領域を抽出することができる。

【００８０】なお、この実施例では、誘目度の評価関数の各係数は、［ｗｈ１，ｗｈ２，ｗｈ３，ｗｈ４］＝［０．０３９，０．０１０，０．０２７，０．０２０］，［ｗｆ１，ｗｆ２，ｗｆ３］＝［０．１３２，０．００５，０．１００］，［ｍ，ｎ］＝［１．３５８，４．２５０］を用い、数１４のθを２０度とし、また視距離ｄをｉｍとし、Ｐ型フーリエ係数を第１０次までとした。

【００８１】また、この実施例では、上記８つの物理的特徴を用いて誘目度を評価するようにしたが、これはあらゆる特徴を有する複数の画像に適合させるためであり、全ての物理的特徴について必ずしも評価する必要はない。

【００８２】続いてステップＳ４１では、決定した注目領域に隣接する図領域を求め、注目領域と隣接図領域との色差およびテクスチャ特徴ベクトルのユークリッド距離を求め、色差が２．０以内であり、かつテクスチャ特徴ベクトルのユークリッド距離が０．３以内であるものを注目領域とともに抽出して、処理を終了する。なお、色差は、上述したように、数５に示すようなＣＩＥ Ｌ＊ａ＊ｂ＊知覚均等色空間における色差式により求められる。また、テクスチャ特徴ベクトルは、数６に従って求められる。

【００８３】図１に戻って、注目領域抽出装置２０で抽出された注目領域は、構図切り取り装置２２に与えられる。構図切り取り装置２２にはメモリ２４が接続され、メモリ２４には画家が描いた絵や写真家が撮影した写真に関する複数のデータが記憶されている。構図切り取り装置２２は、メモリ２４に記憶されたデータを参照して、画家や写真家の構図に合わせて原画像から注目領域を切り取る。具体的には、構図切り取り装置２２は、図７に示すフロー図に従って注目領域を切り取る。つまり、注目領域抽出装置２０で注目領域および隣接する図領域が抽出されると構図切り取り装置２２は処理を開始し、ステップＳ５１で、注目領域の外周画素を求める。つまり、抽出した注目領域の縁を求める。続くステップＳ５３では、上述の上坂が記載した論文の手法により、第１０次までのＰ型フーリエ係数を求め、注目領域を形状ベクトルにする。

【００８４】ここで、メモリ２４に記憶されたデータは、絵画画像や写真画像などの基準画像に対応する画像データ、その絵画画像や写真画像から被写体を抽出してＰ型フーリエ係数により被写体の外周形状を記述した形状ベクトルのデータ、および図８に示すような被写体の位置情報に対応する位置データである。つまり、被写体の位置データは、基準画像の基準構図のデータであり、被写体を囲む外接矩形の横の辺の長さをａ、縦の辺の長さをｂ、外接矩形の原点（画像左上の頂点に最も近い頂点）の位置を（ｗ１，ｈ１）、外接矩形の終点（画像右下の頂点に最も近い頂点）の位置（画像右下の頂点を基準とした位置）を（ｗ２，ｈ２）とした場合に、数１８のようなベクトルで表される。

【００８５】

【数１８】位置データ［ｋｘ１，ｋｙ１，ｋｘ２，ｋｙ２］＝［ｗ１／ａ，ｈ１／ｂ，ｗ２／ａ，ｈ２／ｂ］

続いて、ステップＳ５５でステップＳ５３で求めた形状ベクトルとメモリ２４に格納された絵画画像や写真画像

の形状ベクトルとのユーグリッド距離を求め、つまり基準構図のデータとのマッチングを実行し、ステップＳ５７で領域の形状が最も類似したデータを取得する。言い換えると、最もユークリッド距離が小さい絵画画像の被写体の位置データを取得する。つまり、複数の画像データから注目領域および隣接する図領域に最適な画像データが選択される。そして、ステップＳ５９で、注目領域の外接矩形を求め、外接矩形の原点と終点から数１９に従って原画像の一部を切り取る。なお、数１９では、＊は乗算を意味する。

【００８６】

【数１９】Ｘ１＝ｘ１－Ｗ＊ｋｘ１

Ｙ１＝ｙ１－Ｈ＊ｋｙ１

Ｘ２＝ｘ２＋Ｗ＊ｋｘ２

Ｙ２＝ｙ２＋Ｗ＊ｋｙ２

ただし、注目領域の外接矩形の幅をＷ、高さをＨ、外接矩形の原点を（ｘ１，ｙ１）、終点を（ｘ２，ｙ２）とする。また、Ｘ１＜１のときＸ１＝１，Ｙ１＜１のときＹ１＝１，Ｘ２＞原画像の幅のときＸ２＝原画像の幅，Ｙ２＞原画像の高さのときＹ２＝原画像の高さとする。このとき、切り取る矩形（画像）の頂点座標は数２０のように示される。

【００８７】

【数２０】切り取る画像の頂点座標＝［（Ｘ１，Ｙ１），（Ｘ１，Ｙ２），（Ｘ２，Ｙ１），（Ｘ２，Ｙ２）］

続いて、ステップＳ６１で切り取った結果（画像）を出力して、処理を終了する。

【００８８】したがって、図９（Ａ）に示すような絵画画像の構図に合わせて、図９（Ｂ）に示すような注目領域および隣接する図領域を切り取ることができる。

【００８９】この実施例によれば、上述したような物理的特徴に従って原画像から最も目立つ領域（注目領域）を抽出するので、人間の主観に適合した注目領域を抽出することができる。

【００９０】また、抽出した注目領域を画家が描いた絵画や写真家が撮影した写真の構図に合わせて切り取るので、被写体をあたかもその構図で撮影したかのような写真を生成することができる。つまり、バランスのよい構図を自動で決定することができる。

【００９１】なお、この注目領域抽出装置は画像領域の誘目度を人間の主観に適合して求めることができるので、たとえば、ディジタル圧縮画像・映像の品質評価において、評価対象となる画像や映像の各領域の目立ち度（誘目度）に応じて、その領域に重みづけを行うような客観的な評価を実行するような装置に適用することができる。

【００９２】また、印刷分野では、品質管理の自動化において、印刷ずれ等の問題を軽視できる領域とそうでない領域とを誘目度に応じて自動で判断できる装置に適用することができる。

【００９３】さらに、デザイン分野、特に広告に使用するポスター制作等においては、企業が最も訴えたい部分が目立っているのか否かを客観的に評価する装置に適用することができる。

【００９４】また、ユーザが目立つ服装である場合には、注目領域抽出装置を用いてカメラの回転，チルトおよびズームを制御すれば、カメラが自動でユーザを追うことができる。したがって、たとえば所定のタイミングでシャッタを切るようにすれば、ユーザのスナップ写真を撮ることができる。さらに、撮影した画像（原画像）から構図決定装置を用いて自動で構図を決定することにより、バランスの良いユーザのスナップ写真を作成することができる。

【図面の簡単な説明】

【図１】この発明の一実施例を示す図解図である。

【図２】図１実施例に示す画像生成装置の処理の一部を示すフロー図である。

【図３】図１実施例の画像接合装置で原画像を合成する方法を示す図解図である。

【図４】図１実施例に示す注目領域抽出装置の処理の一部を示すフロー図である。

【図５】図４に示すフロー図に従って注目領域を抽出する方法を示す図解図である。

【図６】ガボールフィルタを示す図解図である。

【図７】図１実施例に示す構図切り取り装置の処理の一部を示すフロー図である。

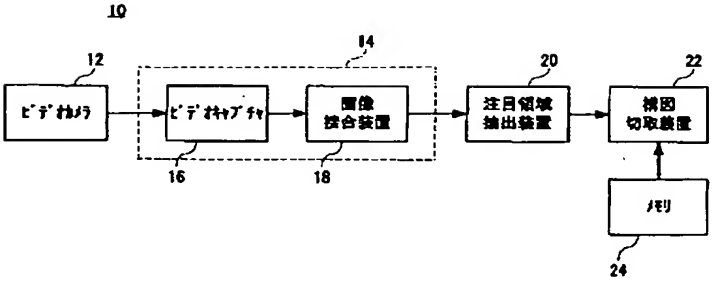
【図８】図７に示すフロー図に従って画像を切り取る場合の位置情報を示す図解図である。

【図９】図７に示すフロー図に従って画像を切り取る場合に参照する画家の絵画画像および参照して切り取った画像を示す図解図である。

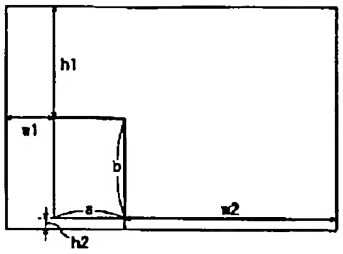
【符号の説明】

１０　…注目領域抽出装置を用いた自動構図決定装置
１２　…カメラ
１４　…画像生成装置
１６　…ビデオキャプチャ
１８　…画像接合装置
２０　…注目度抽出装置
２２　…構図切り取り装置
２４　…メモリ

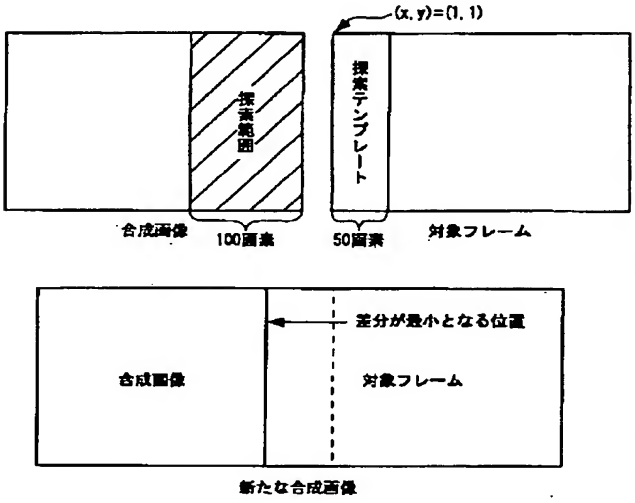
【図１】


10
12
14
16
18
20
22
24
ﾋﾞﾃﾞｵｶﾒﾗ
ﾋﾞﾃﾞｵｷｬﾌﾟﾁｬ
画像
接合装置
注目領域
抽出装置
視図
切取装置
ﾒﾓﾘ

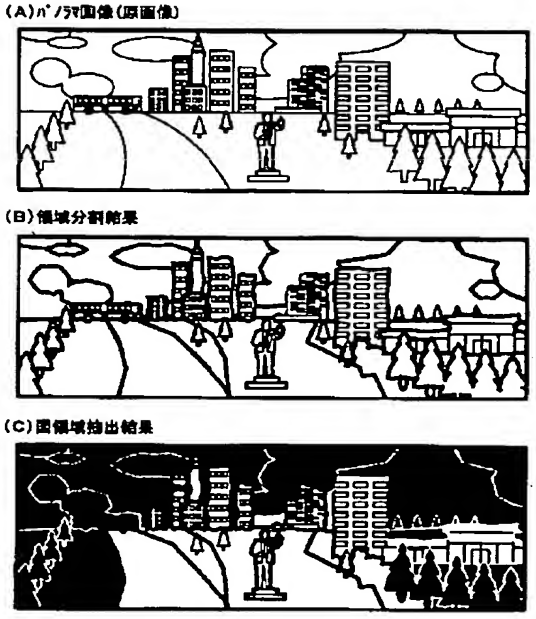
【図８】


h1
w1
b
a
w2
h2

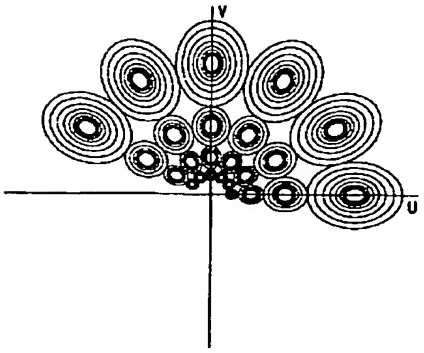
【図３】


(x, y)=(1, 1)
探索範囲
探索テンプレート
合成画像
100画素
50画素
対象フレーム
差分が最小となる位置
合成画像
対象フレーム
新たな合成画像

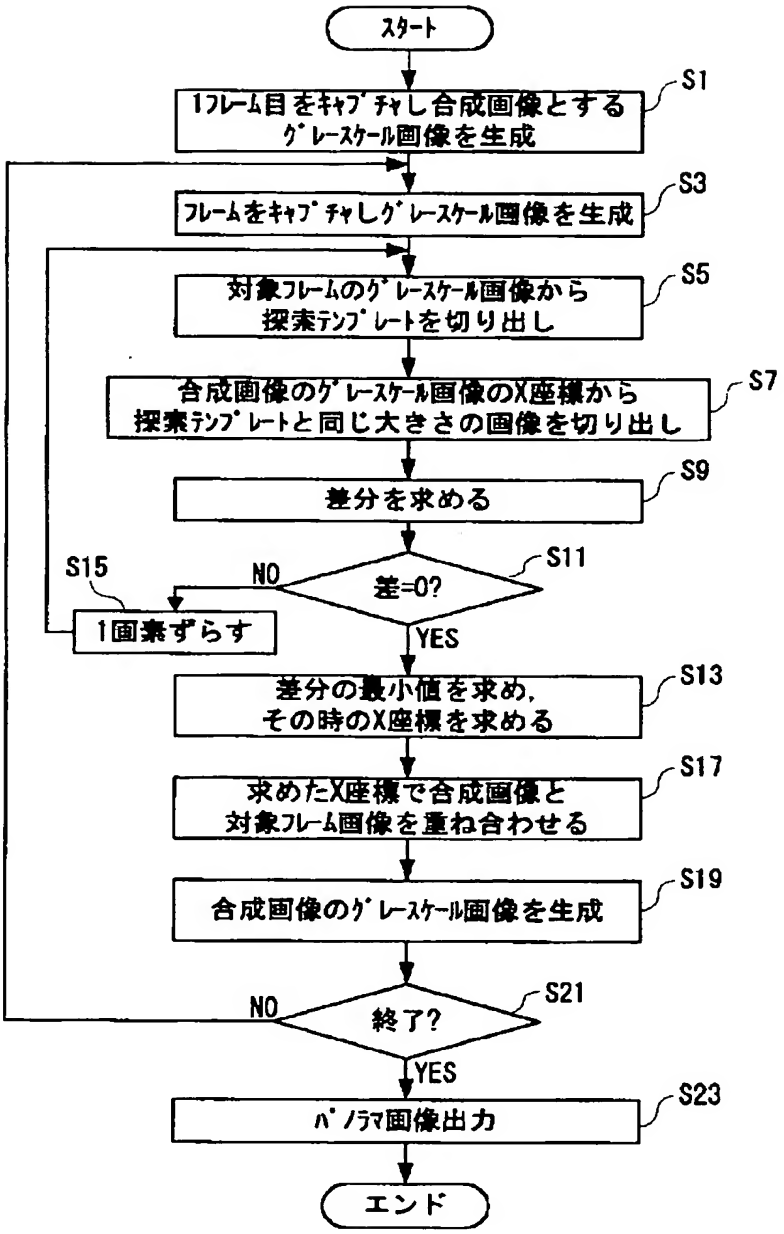
【図４】


(A)ﾊﾟﾉﾗﾏ画像(原画像)
(B)領域分割結果
(C)固領域抽出結果
(D)注目領域抽出結果

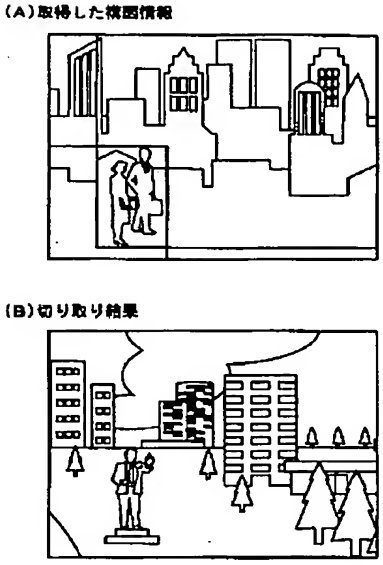
【図６】


V
U

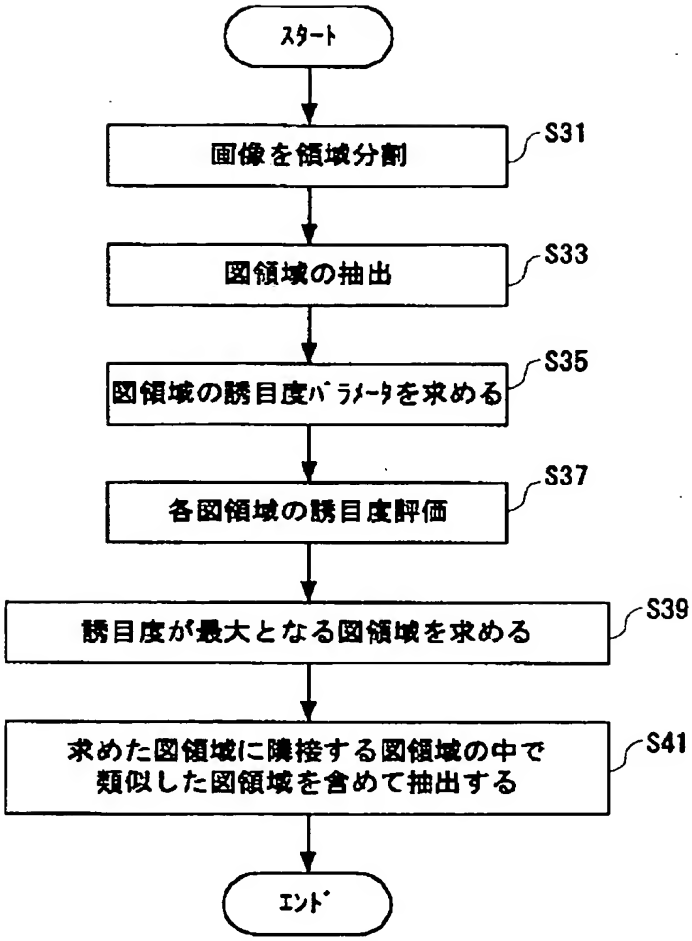
【図２】


スタート
1フレーム目をキャプチャし合成画像とするグレースケール画像を生成
S1
フレームをキャプチャしグレースケール画像を生成
S3
対象フレームのグレースケール画像から探索テンプレートを切り出し
S5
合成画像のグレースケール画像のX座標から探索テンプレートと同じ大きさの画像を切り出し
S7
差分を求める
S9
差=0?
S11
NO
YES
S15
1画素ずらす
差分の最小値を求め，その時のX座標を求める
S13
求めたX座標で合成画像と対象フレーム画像を重ね合わせる
S17
合成画像のグレースケール画像を生成
S19
終了?
S21
NO
YES
パノラマ画像出力
S23
エンド

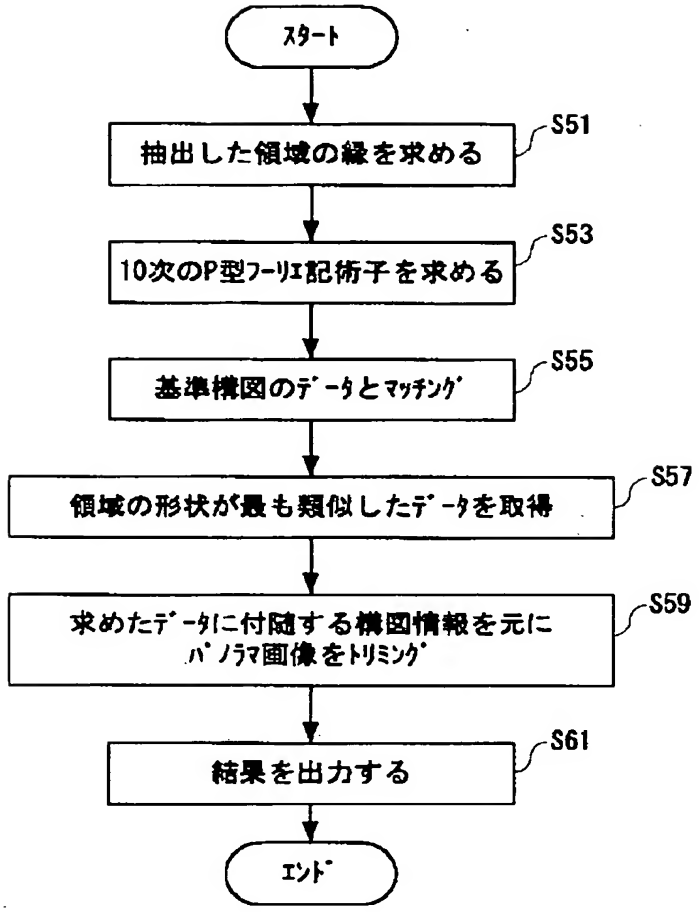
【図９】


(A)取得した複数情報
(B)切り取り結果

【図５】


ｽﾀｰﾄ
画像を領域分割
S31
図領域の抽出
S33
図領域の誘目度ﾊﾟﾗﾒｰﾀを求める
S35
各図領域の誘目度評価
S37
誘目度が最大となる図領域を求める
S39
求めた図領域に隣接する図領域の中で
類似した図領域を含めて抽出する
S41
ｴﾝﾄﾞ

【図７】



フロントページの続き

(51)Int.Cl.'	識別記号	ＦＩ		テーマコード（参考）
		Ｇ０６Ｆ	15/66	４７０Ａ
			15/70	３３０Ｑ
				４６０Ｚ
				４６０Ｂ

(72)発明者　岩舘　祐一
京都府相楽郡精華町大字乾谷小字三平谷５
番地　株式会社エイ・ティ・アール知能映
像通信研究所内

Ｆターム(参考) 5B057 BA15 CA01 CB01 CE09 CE10
CH09 DA08 DB02 DB06 DC04
DC09 DC25
5C022 AA01 AB62 AC27
5C023 AA06 AA16 BA02 BA12 BA13
CA03 DA04
5C076 AA02 AA19 BA03 CA02 CA10
5L096 CA04 FA02 FA15 FA23 FA33
FA46 FA59 FA70 GA38 GA55
JA11

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-126070
(43)Date of publication of application : 11.05.2001

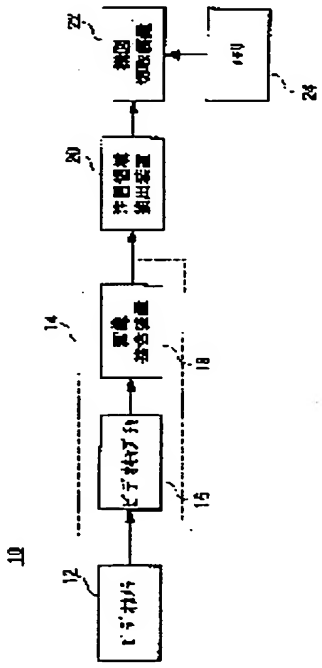
(51)Int.CI. G06T 7/00
G06T 1/00
H04N 1/387
H04N 5/232
H04N 5/262

(21)Application number : 11-306884
(22)Date of filing : 28.10.1999

(71)Applicant : ATR MEDIA INTEGRATION & COMMUNICATIONS RES LAB
(72)Inventor : TANAKA SHOJI
IWADATE YUICHI

(54) NOTICING AREA EXTRACTING DEVICE AND AUTOMATIC COMPOSITION DECIDING DEVICE

(57)Abstract:
PROBLEM TO BE SOLVED: To extract a noticing area adjusted to human subjectively and to automatically device well-balanced composition.
SOLUTION: A noticing area extracting device and an automatic composition deciding device 10 include an image forming device 14, which generates the original image of a panoramic image from video photographed by a camera 12. A noticing area extracting device 20 extracts a noticing area from the original image given from the device 14. Namely, evaluation matched with human subjectivity is given in accordance with the physical feature of the original image and the noticing area is extracted in accordance with the evaluated result. A composition cutting off device 22 cuts out the extracted noticing area and an adjacent image area from the original image by referring to data on paintings painted by painters and photographs taken by photographers stored in a memory 24. Namely, the data can be cut out by the same composition as the case with painted images or photographed images.



LEGAL STATUS

[Date of request for examination] 17.01.2001
[Date of sending the examiner's decision of rejection]
[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]
[Date of final disposal for application]
[Patent number]

[Date of registration]
[Number of appeal against examiner's decision of rejection]
[Date of requesting appeal against examiner's decision of rejection]
[Date of extinction of right]

* NOTICES *

1.This document has been translated by computer. So the translation may not reflect the original precisely.
2.**** shows the word which can not be translated.
3.In the drawings, any words are not translated.

CLAIMS

[Claim(s)]
[Claim 1] An attention field extractor equipped with an evaluation means to be the attention field extractor which extracts an attention field from a subject-copy image, and to evaluate the degree of [**] based on the physical feature, and an extraction means to extract the aforementioned attention field according to the evaluation result of the aforementioned evaluation means.
[Claim 2] The aforementioned physical feature is an attention field extractor containing the degree of different kind of a color according to claim 1.
[Claim 3] For the aforementioned evaluation means, the aforementioned physical feature is an attention field extractor according to claim 2 by which the degree of [******] is evaluated based on at least one degree of different kind among the four degrees of different kind, including further the degree of different kind of the formal degree of different kind, and area, and the degree of different kind of a texture.
[Claim 4] The aforementioned physical feature is an attention field extractor according to claim 1 to 3 which contains a color further.
[Claim 5] The aforementioned physical feature is an attention field extractor according to claim 4 which contains further the area and spatial frequency of a field in the aforementioned subject-copy image.
[Claim 6] The attention field extractor according to claim 1 to 5 further equipped with a photography means to photo a desired image, and a picture generation means to generate the aforementioned subject-copy image based on the aforementioned image.
[Claim 7] The aforementioned picture generation means is an attention field extractor including a junction means to join the aforementioned image photoed with the aforementioned camera for every frame according to claim 6, including the driving means to which the aforementioned photography means carries out the rotation drive of a camera and the aforementioned camera.
[Claim 8] Automatic composition determination equipment equipped with a maintenance means to hold the data about the criteria picture which is automatic composition determination equipment using the attention field extractor according to claim 1 to 7, and has criteria composition, and the cutoff means which cuts out the picture of an attention field from a subject-copy image with reference to the aforementioned criteria composition.
[Claim 9] The aforementioned maintenance means is automatic composition determination equipment according to claim 8 which holds the data about two or more aforementioned criteria pictures, and is further equipped with a selection means to choose the aforementioned data which suited the picture of the aforementioned attention field.
[Claim 10] The aforementioned data are automatic composition determination equipment according to claim 8 or 9 which contains the image data corresponding to the aforementioned criteria picture, a photographic subject's configuration data, and a photographic subject's position data at least.

[Translation done.]

* NOTICES *

1.This document has been translated by computer. So the translation may not reflect the original precisely.
2.**** shows the word which can not be translated.
3.In the drawings, any words are not translated.

DETAILED DESCRIPTION

[Detailed Description of the Invention]
[0001]
[Industrial Application] This invention relates to the attention field extractor cut out from a subject-copy image so that the composition of a picture that experts, such as a painter and a photographer, made the attention field which extracted and extracted the attention field may be suited and it may be especially settled, for example from a subject-copy image about the automatic composition determination equipment which used an attention field extractor and it, and the automatic composition determination equipment using it.
[0002] Here, the field where people observe an attention field in a picture or an image is said.
[0003]
[Description of the Prior Art] As this kind of conventional technology, various methods are proposed about the technique to which a watcher extracts an attention field from a picture. (1) Milanese and Itti And Koch and others assumes the discontinuous part in a picture to be an attention field, and generates the feature map (shade picture) corresponding to two or more physical features acquired from a picture, next asks for the discontinuous part of each feature map, and is extracting as an attention field what unified them.
[0004] (2) Milanese and others filtered the feature map using two or more Difference-of-ofiented-Gaussians filters with which the size was fixed in the discontinuous part of each feature map, and chose and searched for the filtering result from which an output serves as the maximum. Itti, Koch and others normalized each feature map with the square error with each average, and unified all the feature maps by linear combination. And the integrated feature map was recursively filtered with the Difference of Gaussian filter, and the local peak of the filtering result finally obtained was extracted as an attention field.
[0005] Thus, Milanese, Itti and others were extracting the attention field by processing of pixel level, such as filtering and relaxation.
[0006] (3) Martin, Takeuchi and others evaluated the brightness information obtained from a picture based on the information theory of Shannon, and made the attention field the portion with the high amount of information acquired as a result. By this method, the field, i.e., the field which looks complicated, where distribution is large and the bright field of a brightness value are mainly extracted.
[0007] Moreover, with the conventional camera, human being had determined the composition of a photographic subject and a photographic subject manually.
[0008]
[Problem(s) to be Solved by the Invention] However, since the size of an attention field changed with pictures by the technique of of (1) and (2), it was difficult to extract an attention field exactly using the filter with which the size was fixed. Moreover, in the case of the picture which cannot not necessarily be said that the field which looks complicated, or a bright field is in agreement with an attention field like the black picture of a flower vase placed, for example in front of the wall of a complicated pattern by the technique of (3), it was difficult to extract an attention field exactly. Furthermore, by such proposal, there were few examples which performed collating with a watcher's (human being) subjectivity and an extraction result, and it was a

question whether the attention field which actually suits human being's subjectivity can be extracted.
[0009] Moreover, with the conventional camera, when ordinary men without the sense over a photograph determine composition, the good photograph of balance cannot necessarily be taken.
[0010] So, the main purpose of this invention is offering the attention field extractor which can extract exactly the attention field which suited human being's subjectivity.
[0011] Moreover, other purposes of this invention are offering the automatic composition determination equipment which is automatic and can determine the good composition of balance.

[0012]
[Means for Solving the Problem] The 1st invention is an attention field extractor which extracts an attention field from a subject-copy image, and is an attention field extractor equipped with an evaluation means to evaluate the degree of [**] based on the physical feature (whenever for it to be conspicuous), and an extraction means to extract an attention field according to the evaluation result of an evaluation means.
[0013] The 2nd invention is automatic composition determination equipment which used the attention field extractor of a publication for the 1st invention, and is automatic composition determination equipment equipped with a maintenance means to hold the data about the criteria picture which has criteria composition, and the cutoff means which cuts out the picture of an attention field from a subject-copy image with reference to criteria composition.
[0014]
[Function] An evaluation means evaluates the degree of [**] by the attention field extractor of the 1st invention according to the physical feature of a subject-copy image. Here, the degree of [**] means the parameter suitable for human being's subjectivity. An extraction means extracts a most conspicuous field from an evaluation result as an attention field. That is, since an evaluation means carries out evaluation which suited human being's subjectivity according to the physical feature, the attention field which suited human being's subjectivity can be extracted.
[0015] For example, when the physical feature contains the degree of different kind of a color, the degree of [**] can be evaluated based on the difference in the color of each field.
[0016] Moreover, since it contains further the degree of different kind of the formal degree of different kind, and area, and the degree of different kind of a texture (pattern) whenever the color of the physical feature is heterogeneous in addition, if the degree of [**] is evaluated based on at least one degree of different kind of these four degrees of different kind, according to the feature of a subject-copy image, the degree of [**] can be evaluated exactly.
[0017] Moreover, if it is the case where it evaluates also about three elements (a hue, saturation, lightness) of a color, the field near the conspicuous color (red) by human being's subjectivity can be estimated as a most conspicuous field.
[0018] Furthermore, if it evaluates also about the area of each field in spatial frequency or a subject-copy image, most conspicuous evaluation of a field can be judged still more exactly.
[0019] Moreover, the picture corresponding to the image which photoed the desired image with the photography means, for example, was photoed can be compounded, and a subject-copy image can also be generated.
[0020] For example, if the position and height of a camera which are contained in a photography means are fixed, 360 degrees rotates and it enables it to take a photograph, when a junction means joins the picture in every frame, the subject-copy image of a panorama picture is generable within the limits of 360 degrees.
[0021] With the automatic composition determination equipment of the 2nd invention, a maintenance means makes a criteria picture the pictures picture and photograph corresponding to the photograph which the pictures which the painter drew, and the photographer took, and the data about this criteria picture are held. Since a cutoff means cuts out the picture of an attention field from a subject-copy image with reference to the data about this criteria picture, it is automatic and can determine the good composition of balance.
[0022] Moreover, if the data about two or more criteria pictures are held, since a selection means can choose the composition suitable for the picture of an attention field, the good

composition of balance can be determined about all pictures.
[0023] Since they contain the image data corresponding to a criteria picture, a photographic subject's configuration data, and a photographic subject's position data at least, the data about the above criteria pictures choose the composition suitable for the attention field, are automatic and can determine the good composition of balance. That is, a photograph as if it photoed the photographic subject in the composition is generable.
[0024]
[Effect of the Invention] Since an attention field is extracted according to the evaluation result of the physical feature suitable for human being's subjectivity according to this invention, the attention field which suited human being's subjectivity can be extracted.
[0025] According to other invention, since an attention field is cut out from a subject-copy image using the criteria composition of a criteria picture, it is automatic and the good composition of balance can be determined.
[0026] The above-mentioned purpose of this invention, the other purposes, the feature, and an advantage will become still clearer from the detailed explanation of the following examples given with reference to a drawing.
[0027]
[Example] With reference to drawing 1 , the automatic composition determination equipment 10 of this example contains a video camera (only henceforth a "camera") 12. A photography position (a position and height) is fixed using a tripod etc., 360 degrees of cameras 12 rotate and they can be photoed. The image photoed with the camera 12 is inputted into picture generation equipment 14, and the panorama picture as a subject-copy image is generated from the photoed image. In addition, since 360 degrees of cameras 12 can be rotated, picture generation equipment 14 can generate a panorama picture within the limits of 360 degrees.
[0028] Picture generation equipment 14 generates a panorama picture according to the flow view shown in drawing 2 including a video capture 16 and picture junction equipment 18. That is, if photography is started with a camera 12, after picture generation equipment 14 starts processing, is Step S1, carries out the capture of the first video frame (the 1st frame) using a video capture 16 and generates a synthetic picture from the picture corresponding to oneth of them, it will generate a gray-scale picture from the synthetic picture using picture junction equipment 18.
[0029] At continuing Step S3, the capture of the following frame is carried out and a gray-scale picture is generated like the 1st frame. At Step S5, in order to determine from which position the synthetic picture corresponding to the present frame (object frame) should be compounded, a search template is started from the gray-scale picture of an object frame as shown in drawing 3 . The width of face of this search template is 50 pixels, and the height is the same as a frame picture. Moreover, this search plate is cut down from the zero of an object frame.
[0030] Then, at Step S7, the search range is set up from the gray-scale picture of a synthetic picture, and the picture of the same size as a search template is started from the search range. That is, the width of face of the search range is 100 pixels, and the height is the same as a frame picture. Moreover, when the upper left of a frame picture is made into a zero (x y) (= (1 1)), the x-coordinate of a search plate is the position which subtracted only 100 from the width of face of a frame picture, and the y-coordinate is 1.
[0031] Next, in step S9, from the search range, the picture of the same size as a search template is started, and the absolute value of the difference of the mutual pixel value corresponding to the picture and search plate which were cut down is computed. At continuing Step S11, it judges whether the difference of a pixel value is minimum (0). If it is "YES" at Step S11, the started picture will judge that it is the same size as a search plate, and will progress to Step S13. On the other hand, if it is "NO" at Step S11, after judging that the sizes of the picture and search plate which were cut down differ and shifting the search range of 1 pixel to the positive direction of a x axis at Step S15, it will return to Step S5. Thus, processing is repeated until the picture and search template which it started 1 pixel at a time with staggering become the same size.
[0032] At Step S13, an X coordinate in case the absolute value of a difference is minimum (0) is

computed. At continuing Step S17, a synthetic picture and an object frame picture are piled up by the computed X coordinate, and a new synthetic picture is generated. And at Step S19, it judges whether the gray-scale picture was generated from the generated synthetic picture, and the panorama picture was generated at Step S21. If the panorama picture will not be generated if it is "NO" at Step S21 that is, it returns to Step S3. On the other hand, if it is "YES" at Step S21, it will judge that the panorama picture (subject-copy image) as shown in drawing 4 (A) was generated, a subject-copy image will be outputted to the latter attention field extractor 20 (drawing 1) at Step S23, and processing will be ended.

[0033] Returning to drawing 1 , the attention field extractor 20 extracts a most conspicuous field (attention field) from the given subject-copy image. Specifically, the attention field extractor 20 processes according to the flow view shown in drawing 5 . That is, if a subject-copy image is given from picture generation equipment 14, the attention field extractor 20 will start processing and will carry out field division of the subject-copy image at Step S31. Specifically, as shown in drawing 4 (B), a subject-copy image is divided into a drawing field and a picture field. the method of this field division -- 1997IEEE -- setting -- W.Y.Ma and B.S.Manjunath ** -- the boundary detection method based on "edge flow" indicated to "Edge Flow:A Framework of Boundary Detection and Image Segmentation" is applied If it explains briefly, this method is edge flow which asks for the change direction of a color or a pattern in each place of a picture, and consists of strength (size) of the change direction and change. A vector is determined. And it is edge flow by repetitive operation. A vector is spread in the direction of each vector, and let the place where final vectors collide be the boundary line of each field.

[0034] Therefore, the divided drawing field as shown in drawing 4 (C) is extracted, and Step S33 estimates the degree of a drawing field of [**] at Step S35. That is, it asks for the degree parameter of a drawing field of [**]. Here, the subjectivity evaluation experiment which the artificer conducted showed that the physical features required for evaluation of the degree of [**] were the degree of different kind of the degree of different kind of the degree of different kind of a color, and a texture, and a form and the degree of different kind of area, a color, spatial frequency, and area. Moreover, generally, since the relation between human being's subjectivity evaluation result and a physical characteristic is expressed with a serpentine curve in many cases, an one ****** evening function is used for evaluation which is the degree of [**] several 1.

[0035]

[Equation 1]

$$H(m,n)=\int_0^1 h^{m-1}(1-h^{n-1})\,dh$$

[0036] The performance index of the degree of [**] as shown in several 2 is defined using this beta function.

[0037]

[Equation 2]

$$A_{i,t}=\int_0^1 a_{i,t}^{m-1}(1-a_{i,t}^{n-1})\,da_{i,t}$$

$$a_i=HET_i+FP_{i,t}$$

ただし、

$A_{i,t}$: 領域iの時間tにおける誘目度
$a_{i,t}$: 領域iの時間tにおける特徴量統合値
HET_i : 領域iの異質性誘目度
$FP_{i,t}$: 領域iの時間tにおける特徴誘目度
i : 領域i
t : 時間
m, n : 立ち上がり, 飽和パラメータ

[0038] Moreover, the degree HET to the above-mentioned four degrees of different kind of each field shown in several 2 of [**] is defined by several 3.
[0039]
[Equation 3]
$HET_i = wh_1 \cdot HC_i + wh_2 \cdot HT_i + wh_3 \cdot HS_i + wh_4 \cdot HSh_i$

ただし、

HET_i : 領域iの異質性誘目度
HC_i : 領域iの色の異質度
HT_i : 領域iのﾃｸｽﾁｬの異質度
HS_i : 領域iの面積の異質度
HSh_i : 領域iの形の異質度
wh_i : 重み係数

[0040] Furthermore, the degree H of different kind of each physical feature is computed according to several 4, when standard deviation of dm and Difference d is set [the difference of the feature value and the average feature value of all fields] to std for the average of d and Difference d.
[0041]
[Equation 4]
$$H(d, dm, std) = \left| \frac{d - dm}{std} \right|$$

[0042] That is, the degree HC of different kind of the color of each field is CIE. L* a* b* The result which computed and computed the standard deviation of the color difference of the average color of a field and the average color of all fields, a color difference average, and the color difference using the color difference formula in a consciousness equal color space is substituted for several 4, and it asks. In addition, the color difference formula is indicated in detail by 1994 "a color science handbook." Specifically, it is the sexual desire news (R, G, B) of each pixel of a picture Equal perceived-color-space L* a* b* It changes and Euclidean distance in the space is made into the color difference. That is, a color difference formula is shown like several 5.
[0043]
[Equation 5]
$$\Delta Eab^* = [(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)]^{\frac{1}{2}}$$

[0044] Next, the degree HT of different kind of a texture is explained. a texture -- 1996IEEE -- setting -- B.S.Manjunath W.Y.Ma ** -- it is expressed by the texture feature vector indicated to "Texture Features for Browsing and Retrieval of Image Data", and the difference between textures is expressed by the Euclidean distance between texture feature vectors
[0045] Here, a texture feature vector is expressed by the vector which uses as an element the response when filtering a picture by the Gabor filter bank constituted from two or more Gabor filters with which a size differs from a direction. However, since there is no orthogonality in each response when it filters with two or more Gabor filters, redundant information may be included in the filtered result. Then, the parameter of each filter in the Gabor filter bank is determined by the technique indicated by above-mentioned 1996IEEE. Specifically, it asks for theta, a, sigmau, and sigmav of the Gabor filter shown by several 6 by several 7. In addition, for this technique, the filter which adjoins as shown in drawing 6 is Half-Peak. In order to determine that the scale (size) and direction parameter of a filter touch and to express a textural facility, 24 filters (four scales, six directions) are used.
[0046]
[Equation 6]

$G_{mn}(x, y)=a^{-m}\cdot G(x', y')$

$$G(x, y)=\left(\frac{1}{2\pi\sigma_x\sigma_y}\right)\exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)\right]\cdot\exp(2\pi jUx)$$

$U=a^{m}$

$x'=x\cos\theta+y\sin\theta$

$y'=-x\sin\theta+y\cos\theta$

$$\sigma_x=\frac{1}{2\pi\sigma_u}$$

$$\sigma_y=\frac{1}{2\pi\sigma_v}$$

ただし、

Gmn(x, y)： ｽｹｰﾙID=m, 方位ID=nの時のｶﾞﾎﾞｰﾙﾌｨﾙﾀ
m ： 0 .. S-1
n ： 0 .. K-1
S ： ｽｹｰﾙ数
K ： 方位数

[0047]
[Equation 7]

$$\theta=\frac{n\pi}{K}$$

$$a=(U_h/U_l)^{\frac{1}{S-1}}$$

$$\sigma_u=\frac{(a-1)U}{(a+1)\sqrt{2\ln 2}}$$

$$\sigma_v=\tan\left(\frac{\pi}{2k}\right)\left[U_h-2\ln 2\left(\frac{\sigma_u^2}{U_h}\right)\right]\left[2\ln 2-\frac{(2\ln 2)^2\sigma_u^2}{U_h^2}\right]^{\frac{1}{2}}$$

ただし、

Uh: ｶﾞﾎﾞｰﾙﾌｨﾙﾀﾊﾞﾝｸの最大中心周波数
Ul: ｶﾞﾎﾞｰﾙﾌｨﾙﾀﾊﾞﾝｸの最小中心周波数

[0048] A texture feature vector is expressed by this several 6. Therefore, the degree HT of different kind of the texture of each field is computed by substituting the average of distance with the mean vector of all fields, and distance, and the standard deviation of distance for several 4.
[0049] Furthermore, the degree HS of different kind of the area of each field is computed by substituting the average of a difference with the average area of all fields, and a difference, and

the standard deviation of a difference for several 4.
[0050] In the degree of different kind of the hole which the degree of different kind of the appearance of a field and a field include, it unifies and the degree HSh of different kind of the form of each field is obtained further again, as shown in several 8.
[0051]
[Equation 8]
$$HSh_i = \frac{SHD_i + HOD_i}{2}$$

ただし、

HSh_i : 領域iの形の異質度
SHD_i : 領域iの外形の異質度
HOD_i : 領域iの穴の異質度

[0052] Here, it is necessary to take into consideration both the difference in the configuration itself, and the difference in the way of being visible depended on rotating a figure about the difference in a form. Then, a configuration can be described using the P type Fourier descriptor which Yoshinori Kamisaka indicated to "the new Fourier descriptor applicable also to an open curve" in 1984/3 "a telecommunication society paper magazine." In this case, if the power to each frequency after the P type Fourier transform is detected, power is in agreement when the configuration of two figures is the same. Moreover, if a Fourier descriptor is used, it restricts to the time when a configuration and an angle of rotation are in agreement, and a Fourier descriptor is in agreement. Therefore, a Fourier descriptor and the power to each frequency can estimate how many configurations and appearance of two figures are in agreement.
[0053] As mentioned above, the feature of an appearance can be expressed by the vector which consists of a Fourier coefficient and power, and the difference of an appearance can be expressed using the Euclidean distance between these feature vectors. Therefore, the degree HSh of different kind of the appearance of each field is computed by substituting the average of distance with the mean vector of all fields, and distance, and the standard deviation of distance for several 4.
[0054] Next, the degree Hho of different kind of a hole is explained. In order to express the feature of the hole which a field includes, it is necessary to take into consideration not only the configuration of a hole but the number and hole site of a hole. Asking for the primary moment of a field can define the difference in a hole site. In addition, the primary moment is indicated by 1992 "an image-analysis handbook." Therefore, the feature of a hole was expressed by the vector as shown in several 9. In addition, turn of a hole is made into the order near the zero of a picture.
[0055]
[Equation 9]
$$ho = [nh, M, ha_{11}, \cdots, ha_{1n_1}, hb_{21}, \cdots, hb_{2n_2}, \cdots, ha_{m1}, \cdots, ha_{mn_1}, hb_{m1}, \cdots, hb_{mn_2}]$$

ただし、

ho : 穴の形の特徴ﾍﾞｸﾄﾙ
nh_i : 領域iの穴の数
M_i : 領域iのﾓｰﾒﾝﾄ
$ha_{i,1,j}$: 領域iの1番目の穴のj次のﾌｰﾘｴ係数
$hb_{i,1,j}$: 領域iの1番目の穴のj次のﾊﾟﾜｰ

[0056] Therefore, the degree Hho of different kind of the hole of each field is computed by substituting the average of distance with the mean vector of all fields, and distance, and the standard deviation of distance for several 4.
[0057] Moreover, the degree FP of [******] shown in several 2 can be defined as shown in several 10.
[0058]
[Equation 10]